目　　录

绪论……………………………………………………………1
第一章　微分方程预备知识……………………………………14
　§1　勒贝格积分……………………………………………14
　§2　勒贝格不定积分………………………………………27
　§3　具可积右端的常微分方程组…………………………32
　§4　解对参数的连续依赖性………………………………37
　§5　线性微分方程组和它的共轭方程组…………………44
第二章　线性系统的最优控制…………………………………50
　§1　引言……………………………………………………50
　§2　凸集……………………………………………………51
　§3　线性系统的等时区域…………………………………63
　§4　最优开关控制器的设计原理…………………………80
　§5　最优开关控制的唯一性………………………………95
　§6　状态变量为线性的系统………………………………98
　§7　最优线性反馈调节器的设计原理……………………104
　§8　一个线性最优控制问题………………………………112
第三章　庞特里雅金的最大值原理……………………………118
　§1　最优控制问题的叙述和最大值原理…………………118
　§2　动态规划方法与最大值原理…………………………124
　§3　最大值原理与古典变分学……………………………131
　§4　线性系统的二次最优控制问题………………………145
第四章　最大值原理的证明……………………………………150
　§1　引言……………………………………………………150

§ 2　轨线的变分……………………………………………………154
§ 3　最大值原理的证明……………………………………………157
§ 4　横截条件…………………………………………………………166
§ 5　时间不固定的最优控制问题………………………………175
§ 6　时滞系统的最优控制………………………………………184
§ 7　艾克兰变分原理与最大值原理的证明…………………192
第五章　最优控制的近似计算方法…………………………202
§ 1　引言………………………………………………………………202
§ 2　解无约束最优控制问题的梯度法………………………203
§ 3　牛顿法与边值问题……………………………………………219
§ 4　解约束最优控制问题的罚函数方法……………………236
第六章　分布参数系统的最优控制…………………………250
§ 1　引言………………………………………………………………250
§ 2　希尔伯特空间上二次泛函的极值问题…………………251
§ 3　椭圆型系统的最优控制问题………………………………264
§ 4　抛物型系统的最优控制问题………………………………271
§ 5　时间最优控制…………………………………………………283
附录 1　巴拿赫空间………………………………………………294
附录 2　变分学基础………………………………………………310
附录 3　布劳维尔不动点定理…………………………………318
附录 4　菲利浦夫引理……………………………………………323
附录 5　波赫纳积分………………………………………………326
附录 6　最大值原理的非光滑分析证明……………………331
参考书籍……………………………………………………………345
参考文献……………………………………………………………347

绪　论

由于自动控制技术、通讯技术的发展和推动，数学家维纳(N. Wiener)等在本世纪四十年代创立了控制论科学．其后，自动控制技术、自动控制理论、数字计算机技术、航天技术迅猛发展，要求人们根据性能指标设计最优的控制系统．在数学家列甫希茨(S. Lefschetz)的影响下，1953年布绍(D. W. Bushaw)讨论了最优开关控制，给出了二阶线性定常系统的时间最优控制问题的完整解答．钱学森在1954年出版的《工程控制论》一书中介绍了布绍的工作；并在另一章指出，变分学是设计最优控制器的数学方法，同时分析了一些具体实例．但是，布绍所讨论的问题却无法用古典变分学方法进行讨论．

最优控制系统的设计问题提出后，吸引了一大批数学家研究它的数学理论．本世纪五十年代，在苏联以数学家庞特里雅金(Л. С. Понтрягин)为首组织了由数学家和自动控制理论专家组成的讨论班，系统地讨论控制过程的数学理论．1956年，庞特里雅金把最优控制过程问题正确地叙述为具有约束的非古典变分学问题，同时以猜想的形式提出了解决该问题的数学方法——最大值原理，并且用它讨论了布绍所研究的问题，证明了最优开关原理，显示了最大值原理在解决最优控制过程问题中的效用．到1960年，庞特里雅金等完成了最大值原理的严格证明．与此同时，1957年，苏联的数学家克拉索夫斯基(Н. Н. Красовский)运用泛函分析的矩量问题方法来讨论布绍问题的推广——线性系统的时间最优控制问题，直接证明了最优开关原理．

本世纪五十年代，在美国兰德(RAND)公司工作的数学家贝

尔曼(R. Bellman)等也讨论了控制过程的数学问题，发展了动态规划方法，并且也推广了布绍的结果. 拉萨尔(J. P. LaSalle)则用向量值测度理论证明了一般形式的开关原理. 卡尔曼(R. E. Kalman)深入研究了线性系统在二次性能指标下的最优控制问题，把它归结为黎卡提(J. F. Riccati)方程的求解，建立了最优线性反馈调节器的设计理论.

1960年，国际自动控制联合会(IFAC)第一届世界大会在莫斯科举行. 贝尔曼、卡尔曼、庞特里雅金等在大会上报告了他们各自的工作，引起了人们极大的重视，宣告了最优控制理论的诞生.

最优控制的必要条件——最大值原理，把古典变分学中极值曲线的必要条件、最优开关原理等做为应用实例. 但是，古典变分学不能处理最优开关控制问题. 所以人们说最优控制理论是变分学适应控制过程问题的新发展. 然而，美国的伯科维茨(L. D. Berkovitz)1961年指出，运用芝加哥(Chicago)学派在本世纪三十年代发展的变分学方法，也可证明最大值原理. 但是，庞特里雅金关于最优控制问题的叙述和最大值原理是与控制系统的最优设计问题紧密结合的，所以人们还是愿意把最优控制理论视为变分学的新发展，而不把它归结为变分学的一部分.

微分方程是描述控制系统的数学工具，所以微分方程理论成为最优控制系统设计理论的基础和工具.

本书讲述最优控制系统的微分方程理论. 主要内容：最优控制过程问题的来源以及数学描述；线性系统的时间最优控制问题，最优开关原理及其证明；最优线性反馈调节器的设计理论；最大值原理及其在变分学和开关原理、调节器设计理论中的应用，最优控制的计算方法；最大值原理的证明；分布参数系统的最优控制问题等. 它是作为控制科学、应用数学、系统科学以及数学系高年级的选修课或研究生的基础课《最优控制理论》的教材而编写的.

本书要求读者具有微积分、线性代数、常微分方程的基础知识．具有大学数学系、应用数学系二年级和高等工业院校数学知识的读者都可以阅读．本书简要介绍了勒贝格积分初步和凸集理论初步，它是完全理解和掌握本书内容的重要的预备知识．

本书也可作为电子工程、计算机科学、自动控制、管理科学、经济科学等的最优控制理论课的教学用书或教学参考书．对于想了解最优控制理论的数学、应用数学、控制科学和系统科学的工作者，工程技术人员，管理工作人员，本书也有参考作用．

为了解控制问题是怎么回事，我们先从几个简单的例子谈起．

例 1 电梯的最快升降问题

设有一部运送货物的电梯．现要把货物从地面运送到高为 H 的楼面，试问电梯应怎样运行才能把货物最快地运送到目的地？

为简化问题，我们假设电梯的动力是由具有一定转矩的可逆电动机提供的，并且可以随遇静止，即电动机不运行时，电梯在任何高度可以静止．同时，假设摩擦是可以忽略的．

设电梯的质量为 m，货物的质量和电梯的质量相比是很小的，电梯离地面的高度为 x，电动机作用于电梯的力为 $F(t)$，那末电梯的运动方程是

$$m\frac{d^2x}{dt^2}=F(t). \tag{1}$$

因为电动机对电梯的作用力是有限制的，所以存在常数 F_0 使得成立不等式

$$|F(t)|\leqslant F_0. \tag{2}$$

由于电梯初始时处于地面且静止，所以有

$$x(0)=0,\quad \dot{x}(0)=0. \tag{3}$$

如果电梯在 t_1 时刻到达高为 H 的楼面，应当停止运行，所以应当有

$$x(t_1)=H,\quad \dot{x}(t_1)=0. \tag{4}$$

我们的控制问题是：设计满足约束条件(2)的电梯的动力$F(t)$，使得方程(1)的满足初值条件(3)的解适合终值条件(4)，并且时间t_1取得最小值.

经验告诉人们，当$x(t)\leqslant H/2$时，应使电梯以最大的加速度运行，因此

$$F(t)=F_0;$$

当$x(t)>H/2$时，电梯应以最大的减速度运行，即

$$F(t)=-F_0.$$

当$x(t)=H$时，$F(t)=0$. 以这种方式来控制电动机，电梯将以最短时间到达目的地. 该经验是否对？应如何证明呢？

该问题是契贝晓夫(П. Л. Чебышев)于十九世纪在讨论起重机运行时提出的，他在该问题和其他机械问题的基础上发展了函数最优逼近理论. 自然，在当时的技术条件下，他不可能系统地提出控制理论问题.

下面介绍契贝晓夫解决这个问题的方法.

根据(1)、(3)和(4)，控制作用力$F(t)$应满足下述等式

$$\int_0^{t_1}F(t)dt=m\left\{\frac{dx(t_1)}{dt}-\frac{dx(0)}{dt}\right\}=0,$$

$$\int_0^{t_1}\int_0^{s}F(t)dtds=mx(t_1)=mH.$$

即成立着等式

$$\int_0^{t_1}F(t)dt=0,$$

$$\int_0^{t_1}(t_1-t)F(t)dt=mH.$$

所以

$$\int_0^{t_1}F(t)dt=0,$$

$$\int_0^{t_1}tF(t)dt=-mH.$$

因此，对任何常数c，成立着等式

$$\int_0^{t_1}(t+c)F(t)dt=-mH.$$

从而

$$mH=\left|\int_0^{t_1}(t+c)F(t)dt\right|$$

$$\leqslant\int_0^{t_1}|t+c||F(t)|dt\leqslant F_0\int_0^{t_1}|t+c|dt.$$

因此，我们得到不等式

$$mH/F_0\leqslant\inf_c\int_0^{t_1}|t+c|dt. \tag{5}$$

但是，积分

$$\int_0^{t_1}|t+c|dt=\int_c^{t_1+c}|t|dt$$

在 $c=-t_1/2$ 时达到最小值

$$\int_{-t_1/2}^{t_1/2}|t|dt=2\int_0^{t_1/2}tdt=t_1^2/4.$$

所以，由(5)得到

$$\frac{mH}{F_0}\leqslant\frac{1}{4}t_1^2,$$

即时间 t_1 必须满足不等式

$$t_1\geqslant2\sqrt{\frac{mH}{F_0}}.$$

设

$$t^*=2\sqrt{\frac{mH}{F_0}}.$$

我们要证明：t^* 是最短时间，并且

$$F^*(t)=-F_0\,\mathrm{sgn}(t-t^*/2) \tag{6}$$

能使电梯在时刻 t^* 到达高度为 H 的楼面。

事实上，这时

$$m\frac{d^2x(t)}{dt^2}=-F_0\,\mathrm{sgn}(t-t^*/2),$$

所以，从 0 积分到 t^*，得到

$$m\frac{dx(t^*)}{dt}=0,$$

$$\begin{aligned}mx(t^*)&=-\int_0^{t^*}(t^*-t)F_0\operatorname{sgn}(t-t^*/2)dt\\&=\int_{-t^*/2}^{t^*/2}(t-t^*/2)F_0\operatorname{sgn}t\,dt\\&=2F_0\int_0^{t^*/2}t\operatorname{sgn}t\,dt=F_0t^{*2}/4.\end{aligned}$$

因此

$$\dot{x}(t^*)=0,\quad x(t^*)=H.$$

这就是说，在控制作用力(6)的作用下，电梯能以 t^* 时间到达高度 H 的楼面.

根据(6)，我们有

$$F^*(t)=\begin{cases}F_0, & \text{当 } 0\leqslant t\leqslant t^*/2 \text{ 时},\\ -F_0, & \text{当 } t^*/2<t<t^* \text{ 时}.\end{cases}$$

显然，当 $0\leqslant t<t^*/2$ 时

$$mx(t)=\int_0^t(t-s)F_0\,ds=F_0t^2/2,$$

所以

$$x(t)\leqslant H/2;$$

并且

$$x(t^*/2)=H/2.$$

而当 $t>t^*/2$ 时，我们有

$$x(t)>H/2.$$

因此，我们可以说，当 $0\leqslant x\leqslant H/2$ 时，控制作用力 F 等于 F_0，当 $H/2<x<H$ 时，F 等于 $-F_0$，即

$$F=\begin{cases}F_0, & \text{当 } 0\leqslant x\leqslant H/2 \text{ 时},\\ -F_0, & \text{当 } H/2<x<H \text{ 时}.\end{cases}\tag{7}$$

这正是我们凭经验得出的结论.

在控制作用(6)下，方程(1)成为

$$m\frac{d^2x(t)}{dt^2}=-F_0\,\mathrm{sgn}(t-t^*/2), \tag{8}$$

而控制作用(7)下，方程(1)成为

$$m\frac{d^2x}{dt^2}=-F_0\,\mathrm{sgn}(x-H/2). \tag{9}$$

应当注意的是，方程(8)和(9)的右端都是不连续函数．它们的差别是：(8)的右端是时间变量 t 的函数，而(9)的右端是未知变量 x 的函数．方程(9)是通过对(8)的讨论得到的．

在常微分方程的基础课中，我们仅讨论过右端为连续的方程．对于方程(8)，即右端为自变量 t 逐段连续的方程，我们可以利用逐段积分法来解决．对于方程(9)讨论就困难了．

还应指出，为解决电梯最快升降这样简单的控制问题，如果限于 $F(\cdot)$ 是连续函数，将得不到解决．这是我们必须重视的．

人们自然要问：这种方法能否推广到更一般的问题呢？我们建议读者就下述系统

$$\frac{d^nx}{dt^n}=u(t),$$

$$|u(t)|\leqslant 1$$

来讨论类似问题．

本书第二章将讨论该问题的推广，并给出它的一般解法．

例2 空间飞行器软着陆月球的问题

设有一飞行器离月球表面的高度为 h，在月球引力的作用下垂直向月球表面降落，试问应怎样控制喷气燃料，以使飞行器软着陆于月球表面，并且燃料最省．

设飞行器(包括喷气燃料)在时刻 t 的质量为 $m(t)$，离月球表面的高度为 $x(t)$，喷气的推力为 $u(t)$，月球的引力加速度为常数 g，那末飞行器的运动方程为

$$m(t)\frac{d^2x(t)}{dt^2}=-m(t)g+u(t). \tag{10}$$

而喷气的推力 $u(t)$ 应当与喷射的燃料 $dm(t)/dt$ 成正比，所以

$$\frac{dm(t)}{dt}=-ku(t), \tag{11}$$

其中 k 为常数，同时 $u(\cdot)$ 受有约束条件

$$0\leqslant u(t)\leqslant\alpha. \tag{12}$$

这里 α 是常数.

根据假设，初值条件为

$$x(0)=h,\ \dot{x}(0)=0,\quad m(0)=M, \tag{13}$$

终值条件为

$$x(t_1)=\dot{x}(t_1)=0. \tag{14}$$

燃料消耗是

$$m(0)-m(t_1)=k\int_0^{t_1}u(t)dt. \tag{15}$$

我们的控制问题是：选取满足不等式(12)的控制作用 $u(\cdot)$，使得(10)、(11)满足初值条件(13)的解适合终值条件(14)，并且使燃料消耗(15)达到最小值.

这是航天技术提出的一个控制问题．现在还不打算解决这个问题．应当说，航天技术中有大量的控制问题.

例3　一个生产计划问题

设某产品在时刻 t 的库存量为 $x(t)$，已知单位时间的需求量为 $r(t)$．设单位时间库存费用为 $bx(t)$，其中 b 为常数.如果单位时间的生产量为 $u(t)$，所需费用为 $C(u(t))$．设单位时间的最大生产能力为 α，试问在整个时间 T 内应如何安排生产计划，可以使得各种费用的总和为最小.

现根据问题来列写它的数学描述.

显然，成立着

$$\frac{dx(t)}{dt}=-r(t)+u(t), \tag{16}$$

$$x(0)=x_0>0.$$

并且计划生产数 $u(\cdot)$ 应满足条件

$$0\leqslant u(t)\leqslant a, \tag{17}$$

同时应保证库存量 $x(t)$ 满足条件

$$0\leqslant x(t). \tag{18}$$

而在时间 T 内总的费用是

$$J(u(\cdot))=\int_0^T\{C(u(t))+bx(t)\}dt. \tag{19}$$

我们的控制问题是：选取满足条件(17)的生产计划数 $u(t)$，使得(16)的解 $x(\cdot)$ 满足条件(18)，并且使总的费用(19)达到最小值.

这是一个非常简单的生产管理中提出的控制问题．它不仅有约束条件加在控制作用 $u(t)$ 上，而且要求系统的未知量 $x(t)$ 满足一个不等式，即 $x(\cdot)$ 也有约束.

例4 捷线问题

这是贝努利(John Bernoulli) 1696年提出的问题，为牛顿(I. Newton)、莱布尼兹(G. W. Leibniz)、洛毕塔(L'Hospital)、贝努利兄弟(James Bernoulli 和 John Bernoulli)所解决．该问题是：设 $P(x_0, y_0)$ 和 $Q(x_1, y_1)$ 是垂直平面中的两个给定点，$y_0>y_1$，联结 P 和 Q 的光滑曲线为 $y=y(x)$，质量为 m 的质点沿该曲线在重力作用下自由降落，如果没有摩擦，试问怎样的曲线能使质点从点 P 最快到达点 Q？这就是贝努利提出的捷线问题．它是典型的古典变分学问题之一.

设 t 是时间变量，曲线的参数方程为

$$x=x(t),\quad y=y(t), \tag{20}$$

它过 P 和 Q,所以

$$\begin{aligned} x(0)&=x_0, \quad y(0)=y_0, \\ x(t_1)&=x_1, \quad y(t_1)=y_1, \end{aligned} \tag{21}$$

这里 t_1 是质点到达 Q 的时间. (20)也可以理解为质点在 t 时刻达到点 $(x(t), y(t))$.

根据能量守恒定律,我们有

$$\frac{1}{2}m\left\{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2\right\}+mgy=mgy_0,$$

这里 g 是引力常数. 所以

$$\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2=2g(y_0-y).$$

设曲线在时刻 t 处的切线和 y 轴的交角为 $\theta(t)$,那末

$$\frac{dx}{dt}\Big/\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}=\sin\theta(t),$$

$$\frac{dy}{dt}\Big/\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}=-\cos\theta(t),$$

这里的负号是因为 y 随着 t 的增加而减小. 因此,

$$\begin{aligned} \frac{dx}{dt}&=\sqrt{2g(y_0-y)}\,\sin\theta(t), \\ \frac{dy}{dt}&=-\sqrt{2g(y_0-y)}\,\cos\theta(t). \end{aligned} \tag{22}$$

显然,选取角度 $\theta(t)$ 相当于选取曲线(20). 从而,捷线问题是:选取 $\theta(t)$,使得(22)以 (x_0, y_0) 为初值的解满足终值条件

$$x(t_1)=x_1, \quad y(t_1)=y_1.$$

并且时间 t_1 取最小值.

如果置 $$u(t)=\sin\theta(t),$$

那末 $$\cos\theta(t)=\sqrt{1-u^2(t)},$$

而 $u(t)$ 满足条件

$$|u(t)|\leqslant 1. \tag{23}$$

这时(22)可改写为

$$\begin{aligned}&\frac{dx}{dt}=\sqrt{2g(y_0-y)}\,u(t),\\&\frac{dy}{dt}=-\sqrt{2g(y_0-y)}\sqrt{1-u^2(t)}.\end{aligned}\tag{24}$$

捷线问题是：选取满足条件(23)的 $u(\cdot)$，使得方程(24)存在满足条件(21)的解，并且时间 t_1 取最小值.

例 5 反馈调节器的设计问题

设受控对象的输入(控制作用)为 $u(\cdot)$，而输出为 $x(\cdot)$，它的动态方程为

$$\frac{dx(t)}{dt}=u(t).\tag{25}$$

显然，如果控制作用 u 为 0，那末它的输出不变. 我们要求设计反馈调节器，使得系统的输出趋于零，并且系统的品质是较好的. 在自动控制理论中，通常要求控制作用 u 是输出 x 的线性函数，即线性反馈，并且用平方积分指标来衡量控制系统的优劣.

为讨论反馈调节器，我们先考虑开环控制问题，即选取控制作用 $u(t)$，使得平方积分

$$\int_0^\infty\{x^2(t)+u^2(t)\}dt\tag{26}$$

达到最小值. 我们在平方积分指标(26)的被积函数中，不仅引进输出 $x(t)$ 而且引进了控制作用 $u(t)$. 因为

$$\int_0^T\{x^2(t)+u^2(t)\}dt=\int_0^T[\{x(t)+u(t)\}^2-2x(t)u(t)]dt,$$

所以

$$\begin{aligned}&\int_0^T\{x^2(t)+u^2(t)\}dt\\&\quad=-\int_0^T 2x(t)\frac{dx(t)}{dt}dt+\int_0^T\{x(t)+u(t)\}^2dt\end{aligned}$$

$$=x^2(0)-x^2(T)+\int_0^T\{x(t)+u(t)\}^2dt$$

$$\geqslant x^2(0)-x^2(T). \tag{27}$$

令 $T\to\infty$，因为要求 $x(T)\to0$，所以

$$\int_0^\infty\{x^2(t)+u^2(t)\}dt\geqslant x^2(0). \tag{28}$$

又，如果

$$x(t)+u(t)\equiv0, \tag{29}$$

那末
$$\frac{dx(t)}{dt}=-x(t),$$

从而
$$x(t)=x(0)e^{-t},$$

并且(27)成为

$$\int_0^T\{x^2(t)+u^2(t)\}dt=x^2(0)(1-e^{-2T}).$$

令 $T\to\infty$，得到

$$\int_0^\infty\{x^2(t)+u^2(t)\}dt=x^2(0).$$

即，当 $u(\cdot)$ 和 $x(\cdot)$ 适合(29)时，积分(26)达到最小值．因此，从开环控制问题得到线性反馈调节器(29)．

上述五个例题来自不同的领域．例1是研究起重机时引进的，它在空间技术的卫星会合问题中有应用；例2是空间技术提出的问题；例3是数量经济学中的问题；例4是变分学创立时的典型例题；例5是自动控制理论中提出的．

例1和例5表明，为研究反馈调节器或闭环控制问题，宜于先研究开环控制问题．它们也表明：开环最优控制问题的解可以转化为闭环控制问题的解．

以上研究还表明，为研究控制问题，需要根据问题来列写描述它的数学方程，首先是对象的动态方程，也称为系统的状态方程；其次，在这些问题中都有供人们选取的自变量 t 的函数，人们称它

们为控制作用，并且它们受一定条件的约束；第三，还有给定的初值条件和终值条件；最后，为衡量系统的优劣或控制作用的好坏，应当有刻划它们的性能指标．这些，是上述五个例题的共性．

最优控制问题是：选取满足约束条件的控制作用，使系统存在满足初值和终值条件的解，并且使系统的性能指标达到最小值．

读者可以找出上述例题中的状态方程，控制作用，约束条件，初值和终值条件，以及性能指标．

我们将逐步精确地叙述最优控制问题的一般提法，并讲述它们的解法．这里只是一些最简单的例题．

习　题

1　试讨论系统

$$\frac{dx(t)}{dt}=x(t)+u(t),\quad x(0)=x_0,$$

在性能指标

$$\int_0^\infty\{3x^2(t)+u^2(t)\}dt$$

取最小值时的最优控制．

2　对于系统(25)，试分析控制作用

$$u(t)=-x_0e^{-t}$$

和

$$u(t)=-x(t)$$

的特性．

3　试讨论把系统

$$\frac{d^nx}{dt^n}=u(t),\quad |u(t)|\leqslant 1,$$

从状态(0, 0, …, 0)最快地移到$(x_1, 0, \cdots, 0)$的最优控制问题，这里x_1给定．

4　试叙述一般形式的最优控制问题．

第一章　微分方程预备知识

本章讲述本书所需要的，有关勒贝格(H. Lebesgue)积分和具可积右端微分方程的一些预备知识．熟悉这方面知识的读者，可以在用到时参阅，不具备这方面知识的读者，可以从这里学到有关内容．

§1　勒贝格积分

勒贝格积分是实变函数论的基本内容．

我们要简要讲述勒贝格积分的定义和勒贝格可积函数的性质．

如果 E 是实轴上的集，并且对于任给正数 ε，存在有限个或可列个开区间 (α_i, β_i)，它的并集含有 E，且 $\sum_j (\beta_j-\alpha_j)<\varepsilon$，那末称 E 是零集．

显然，有限个或可列个零集 E_j 的并集 $\bigcup_j E_j$ 是零集．

如果 φ 映闭区间 $[a, b]$ 到 $\mathbb{R}$，并且存在有限个 t_j 和 $c_j(j=0, 1, \cdots, l)$，使得

$$a=t_0<t_1<\cdots<t_l=b,$$

$$\varphi(t)=c_j, \text{当 } t\in(t_j, t_{j+1}) \text{ 时},$$

就称 φ 是简单函数．我们并不关心 φ 在 t_j 处取怎样的值．φ 在 $[a, b]$ 上的积分定义为

$$\int_a^b \varphi(t)\,dt=\sum_{j=0}^{l-1} c_j(t_{j+1}-t_j).$$

显然，如果 φ 和 ψ 都是 $[a, b]$ 上的简单函数，α 和 β 是实数，那末 $\alpha\varphi+\beta\psi$ 也是 $[a, b]$ 上的简单函数，并且成立着

$$\int_a^b \{\alpha\varphi(t)+\beta\psi(t)\}dt=\alpha\int_a^b \varphi(t)dt+\beta\int_a^b \psi(t)dt.$$

如果 g 是 $[a, b]$ 上的实函数，并且存在简单函数序列 $\{\varphi_k\}$，满足不等式

$$\varphi_1(t)\leqslant\varphi_2(t)\leqslant\cdots\leqslant\varphi_k(t)\leqslant\cdots$$

除 $\{\varphi_k(\cdot)\}$ 的不连续点外成立，并且在 $[a, b]$ 上除零集外，成立着

$$\lim_{k\to\infty}\varphi_k(t)=g(t),$$

还有
$$\lim_{k\to\infty}\int_a^b \varphi_k(t)dt<+\infty,$$

就称 g 为 C_1 类函数，记为

$$g\in C_1.$$

我们规定 $g\in C_1$ 在 $[a, b]$ 上的积分为

$$\int_a^b g(t)\,dt=\lim_{k\to\infty}\int_a^b \varphi_k(t)dt.$$

可以证明：C_1 类函数的积分的上述定义是不依赖于序列 $\{\varphi_k\}$ 的选取的.

如果 $[a, b]$ 上的实函数 f 可以表示为 C_1 类函数之差，即存在 g_1、$g_2\in C_1$，使得

$$f(t)=g_1(t)-g_2(t),$$

则称 f 在 $[a, b]$ 上是勒贝格可积的，并且定义它在 $[a, b]$ 上的积分为

$$\int_a^b f(t)\,dt=\int_a^b g_1(t)dt-\int_a^b g_2(t)dt.$$

显然，如果还有 h_1、$h_2\in C_1$，使得

$$f(t)=h_1(t)-h_2(t),$$

那末
$$h_1(t)-h_2(t)=g_1(t)-g_2(t),$$

从而
$$h_1(t)+g_2(t)=g_1(t)+h_2(t),$$

而 h_1+g_2 和 g_1+h_2 都是 C_1 类函数，所以

$$\int_a^b \{h_1(t)+g_2(t)\}\,dt=\int_a^b \{g_1(t)+h_2(t)\}\,dt,$$

即
$$\int_a^b h_1(t)\,dt+\int_a^b g_2(t)\,dt=\int_a^b g_1(t)\,dt+\int_a^b h_2(t)\,dt,$$

因此
$$\int_a^b h_1(t)dt-\int_a^b h_2(t)dt=\int_a^b g_1(t)dt-\int_a^b g_2(t)\,dt,$$

即勒贝格积分的定义是不依赖于 C_1 类函数 g_1、g_2 的选取的.

容易证明: 如果 f_1、f_2 在 $[a, b]$ 上是勒贝格可积的, α_1、α_2 是实数, 那末 $\alpha_1 f_1+\alpha_2 f_2$ 在 $[a, b]$ 上是勒贝格可积的, 并且

$$\int_a^b \{\alpha_1 f_1(t)+\alpha_2 f_2(t)\}\,dt$$

$$=\alpha_1\int_a^b f_1(t)\,dt+\alpha_2\int_a^b f_2(t)\,dt.$$

根据勒贝格积分的定义, 如果 f 在 $[a, b]$ 上是勒贝格可积的, 那末存在简单函数列 $\{\varphi_k\}$ 和 $\{\psi_k\}$, 使得

$$\varphi_1(t)\leqslant\varphi_2(t)\leqslant\cdots\leqslant\varphi_k(t)\leqslant\cdots,\qquad \lim_{k\to\infty}\varphi_k(t)=g(t),\qquad (1)$$

$$\psi_1(t)\leqslant\psi_2(t)\leqslant\cdots\leqslant\psi_k(t)\leqslant\cdots,\qquad \lim_{k\to\infty}\psi_k(t)=h(t),\qquad (2)$$

$$f(t)=g(t)-h(t).$$

并且
$$\lim_{k\to\infty}\int_a^b \varphi_k(t)dt<+\infty,$$

$$\lim_{k\to\infty}\int_a^b \psi_k(t)\,dt<+\infty.$$

在(1)、(2)中不妨设 $\varphi_1(t)\geqslant 0$, $\psi_1(t)\geqslant 0$. 否则的话, 取

$$\tilde{\varphi}_k(t)=\varphi_k(t)-\{(\varphi_1(t)+\psi_1(t))-|\varphi_1(t)-\psi_1(t)|\}/2,$$

$$\tilde{\psi}_k(t)=\psi_k(t)-\{(\varphi_1(t)+\psi_1(t))-|\varphi_1(t)-\psi_1(t)|\}/2,$$

那末 $\tilde{\varphi}_k$ 和 $\tilde{\psi}_k$ 是简单函数, 并且

$$0\leqslant\tilde{\varphi}_1(t)\leqslant\tilde{\varphi}_2(t)\leqslant\cdots\leqslant\tilde{\varphi}_k(t)\leqslant\cdots,$$

$$0\leqslant\tilde{\psi}_1(t)\leqslant\tilde{\psi}_2(t)\leqslant\cdots\leqslant\tilde{\psi}_k(t)\leqslant\cdots,$$

$$\lim_{k\to\infty}\tilde{\varphi}_k(t)=g(t)-\{(\varphi_1(t)+\psi_1(t))$$
$$-|\varphi_1(t)-\psi_1(t)|\}/2, \qquad (3)$$
$$\lim_{k\to\infty}\tilde{\psi}_k(t)=h(t)-\{(\varphi_1(t)+\psi_1(t))$$
$$-|\varphi_1(t)-\psi_1(t)|\}/2, \qquad (4)$$

而且 $$f(t)=\lim_{k\to\infty}\tilde{\varphi}_k(t)-\lim_{k\to\infty}\tilde{\psi}_k(t).$$

所以，我们可以假设(1)、(2)中的 $\varphi_1(t)$、$\psi_1(t)$ 为非负的.

如果置

$$\gamma_k(t)=\max(\varphi_k(t),\ \psi_k(t))$$
$$=\frac{1}{2}\{|\varphi_k(t)-\psi_k(t)|+\varphi_k(t)+\psi_k(t)|\},$$
$$k=1,\ 2,\ \cdots,$$

那末 $$\gamma_1(t)\leqslant\gamma_2(t)\leqslant\cdots\leqslant\gamma_k(t)\leqslant\cdots,$$

并且 $$\gamma_k(t)\leqslant\varphi_k(t)+\psi_k(t),$$
$$\lim_{k\to\infty}\int_a^b\gamma_k(t)dt<+\infty,$$
$$\lim_{k\to\infty}\gamma_k(t)=\max(g(t),\ h(t)).$$

所以，$\max(g,\ h)$ 是 C_1 类函数.

同样可证：$\min(g,\ h)$ 也是 C_1 类函数.

因为

$$|f(t)|=|g(t)-h(t)|$$
$$=\max(g(t),\ h(t))-\min(g(t),\ h(t)),$$

所以，$|f|$ 在 $[a,\ b]$ 上是勒贝格可积的，并且成立着不等式

$$\int_a^b|f(t)|dt=\int_a^b\max(g(t),\ h(t))dt$$
$$-\int_a^b\min(g(t),\ h(t))dt$$
$$\geqslant\int_a^b g(t)dt-\int_a^b h(t)\,dt=\int_a^b f(t)dt,$$

同样可得

$$\int_a^b |f(t)|dt \geqslant -\int_a^b f(t)dt,$$

从而

$$\left|\int_a^b f(t)dt\right| \leqslant \int_a^b |f(t)|dt.$$

对于任给正数 ε，存在 k_0 使得

$$0 \leqslant \int_a^b g(t)dt - \int_a^b \varphi_{k_0}(t)dt < \frac{\varepsilon}{2},$$

$$0 \leqslant \int_a^b h(t)dt - \int_a^b \psi_{k_0}(t)dt < \frac{\varepsilon}{2}.$$

置

$$\varphi(t) = \varphi_{k_0}(t) - \psi_{k_0}(t),$$

那末 φ 是简单函数，并且

$$\begin{aligned}\int_a^b |f(t) - \varphi(t)|dt &\leqslant \int_a^b |g(t) - \varphi_{k_0}(t)|dt \\ &\quad + \int_a^b |h(t) - \psi_{k_0}(t)|dt < \varepsilon.\end{aligned}$$

因此，我们证明了下述定理.

定理 1 设 f 在 $[a, b]$ 上是勒贝格可积的，那末对任给正数 ε，存在简单函数 φ，使得

$$\int_a^b |f(t) - \varphi(t)|dt < \varepsilon. \tag{5}$$

设 E 是 $[a, b]$ 中的集，定义

$$\chi_E(t) = \begin{cases} 1, & \text{当 } t \in E \text{ 时}, \\ 0, & \text{当 } t \,\overline{\in}\, E \text{ 时}, \end{cases}$$

称 χ_E 为 E 的特征函数. 如果 χ_E 在 $[a, b]$ 上是勒贝格可积的，就称集 E 是勒贝格可测的，简称 E 是可测的，并且定义集合 E 的测度 $\mu(E)$ 为

$$\mu(E) = \int_a^b \chi_E(t)dt.$$

这样，我们从积分定义出发定义了集合的测度. 而勒贝格在引进积分时是先从集合的测度定义开始的.

当
$$\mu(E)=0$$
时，称 E 为零测度集，它就是我们前面引进的零集.

对于可测集 E 上的实函数 f，如果把它的定义域扩充到闭区间 $[a, b]$，使得当 $t\in[a, b]\setminus E$ 时它为 0，并且扩充了的 f 在 $[a, b]$ 上是勒贝格可积的，就称 f 在 E 上是勒贝格可积的，并且它在 E 上的积分为

$$\int_E f(t)dt=\int_a^b f(t)\chi_E(t)dt.$$

今后简称勒贝格可积函数为可积函数.

设 E 是有限个互不相交的可测集 $E_j(j=1, 2, \cdots, l)$ 的并集，

即
$$E=\bigcup_{j=1}^{l} E_j,\ E_j\cap E_k=\varnothing\ (j\neq k),\ E_j \text{ 可测},$$

那末函数
$$f(t)=\sum_{j=1}^{l} c_j\chi_{E_j}(t)$$
是 E 上的可积函数. 这里 c_1、c_2、$\cdots$、c_l 是实常数.

容易知道
$$\int_E \sum_{j=1}^{l} c_j\chi_{E_j}(t)dt=\sum_{j=1}^{l} c_j\mu(E_j).$$

利用定理 1 容易证明

定理 2 设 f 是可测集 E 上的可积函数，$E\subset[a, b]$，那末对于任给正数 ε，存在 $[a, b]$ 上的简单函数 φ，使得

$$\int_E |f(t)-\varphi(t)|dt<\varepsilon. \tag{6}$$

事实上，在 $[a, b]$ 上对于 $f(t)\chi_E(t)$ 应用定理 1，存在 φ，使得

$$\int_a^b |f(t)\chi_E(t)-\varphi(t)|dt<\varepsilon,$$

但是
$$\int_E |f(t)-\varphi(t)|dt+\int_{[a,b]\setminus E}|\varphi(t)|dt$$

$$=\int_a^b |f(t)\chi_E(t)-\varphi(t)|dt.$$

所以定理 2 成立.

对于可测集 E 上取值于 $\mathbb{R}^n$ 的函数 $\boldsymbol{f}$，如果它的每一分量 $f_j(j=1, 2, \cdots, n)$是 E 上的可积实函数，就称 $\boldsymbol{f}$ 在 E 上是勒贝格可积的，并规定积分的每一分量是

$$\int_E f_j(t)dt \quad (j=1, 2, \cdots, n),$$

即 $$\int_E \boldsymbol{f}(t)\,dt=\left(\int_E f_1(t)dt, \int_E f_2(t)dt, \cdots, \int_E f_n(t)dt\right)^T,$$

这里记号"T"表示转置.

显然，成立着

$$\int_E\{\alpha\boldsymbol{f}(t)+\beta\boldsymbol{g}(t)\}dt=\alpha\int_E \boldsymbol{f}(t)\,dt+\beta\int_E \boldsymbol{g}(t)dt,$$

如果 $E=E_1\cup E_2$, $E_1\cap E_2=\varnothing$，那末

$$\int_E \boldsymbol{f}(t)dt=\int_{E_1} \boldsymbol{f}(t)dt+\int_{E_2} \boldsymbol{f}(t)dt.$$

对于 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T\in\mathbb{R}^n$，我们规定 $\boldsymbol{x}$ 的范数为

$$\|\boldsymbol{x}\|=\left\{\sum_{j=1}^n x_j^2\right\}^{\frac{1}{2}}.$$

显然，成立着不等式

$$\left\|\int_E \boldsymbol{f}(t)dt\right\|\leqslant\int_E \|\boldsymbol{f}(t)\|dt.$$

定理 3 设 $\boldsymbol{f}:[a, b]\to\mathbb{R}^n$ 是勒贝格可积的，正数 $\lambda\in(0, 1)$，那末对于任给正数 ε，存在可测集 $E_\lambda(\varepsilon)\subset[a, b]$，使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\cdot\lambda,$$

$$\lambda\int_a^b \boldsymbol{f}(t)\,dt=\int_{E_\lambda(\varepsilon)} \boldsymbol{f}(t)dt+\boldsymbol{\eta}, \tag{7}$$

$$\|\boldsymbol{\eta}\|<\varepsilon.$$

证 我们来逐步证明定理 3.

1° 设 $\boldsymbol{f}(t)\equiv\mathbf{c}$, 那末显然成立着

$$\lambda\int_a^b \mathbf{c}\,dt=\lambda\mathbf{c}(b-a)=\int_a^{a+(b-a)\lambda}\mathbf{c}\,dt,$$

取 $E_\lambda=[a,\ a+\lambda(b-a)]$, 那末

$$\mu(E_\lambda)=(b-a)\cdot\lambda,$$

$$\lambda\int_a^b \mathbf{c}\,dt=\int_{E_\lambda}\mathbf{c}\,dt,$$

即 (7) 成立, 且 $\boldsymbol{\eta}=\mathbf{0}$.

2° 如果 $\boldsymbol{f}$ 是 $[a,\ b]$ 上的简单函数, 即

$$\boldsymbol{f}(t)=\sum_{j=0}^{l-1}\mathbf{c}_j\chi_{[t_j,t_{j+1}]}(t),$$

其中 $a=t_0<t_1<\cdots<t_l=b$, $\mathbf{c}_j\in\mathbb{R}^n(j=0,\ 1,\ \cdots,\ l-1)$, 那末

$$\lambda\int_a^b\boldsymbol{f}(t)\,dt=\lambda\sum_{j=0}^{l-1}\mathbf{c}_j(t_{j+1}-t_j)$$

$$=\sum_{j=0}^{l-1}\int_{t_j}^{t_j+\lambda(t_{j+1}-t_j)}\mathbf{c}_j\,dt.$$

置 $$E_\lambda=\bigcup_{j=0}^{l-1}[t_j,\ t_j+\lambda(t_{j+1}-t_j)],$$

那末 $$\mu(E_\lambda)=(b-a)\cdot\lambda,$$

$$\lambda\int_a^b\boldsymbol{f}(t)\,dt=\int_{E_\lambda}\boldsymbol{f}(t)\,dt.$$

3° 对于可积的 $\boldsymbol{f}$ 和 $\varepsilon>0$, 存在简单函数 $\boldsymbol{\varphi}$, 使得

$$\int_a^b\|\boldsymbol{f}(t)-\boldsymbol{\varphi}(t)\|\,dt<\varepsilon.$$

根据 2°, 对于 $\boldsymbol{\varphi}$, 存在 $E_\lambda(\varepsilon)\subset[a,\ b]$, 使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\cdot\lambda,$$

$$\lambda\int_a^b\boldsymbol{\varphi}(t)\,dt=\int_{E_\lambda(\varepsilon)}\boldsymbol{\varphi}(t)\,dt.$$

因此, 成立着

$$\lambda\int_a^b f(t)dt=\lambda\int_a^b \varphi(t)dt+\lambda\int_a^b\{f(t)-\varphi(t)\}dt$$
$$=\int_{E_\lambda}\varphi(t)dt+\lambda\int_a^b\{f(t)-\varphi(t)\}dt$$
$$=\int_{E_\lambda} f(t)dt+\int_{E_\lambda}\{\varphi(t)-f(t)\}dt$$
$$+\lambda\int_a^b\{f(t)-\varphi(t)\}dt.$$

记上式有端第二和第三项的和为 η，那末

$$\eta=\lambda\int_{[a,b]\setminus E_\lambda}\{f(t)-\varphi(t)\}dt$$
$$+(1-\lambda)\int_{E_\lambda}\{\varphi(t)-f(t)\}dt.$$

从而

$$\|\eta\|\leqslant\lambda\int_{[a,b]\setminus E_\lambda}\|f(t)-\varphi(t)\|dt$$
$$+(1-\lambda)\int_{E_\lambda}\|\varphi(t)-f(t)\|dt<\varepsilon.$$ ∎

定理 4 设对于 $t\in[\alpha,\beta]$，$f(t,\cdot):[a,b]\to\mathbb{R}^n$ 是可积的，并且成立着

$$\lim_{\bar t\to t}\int_a^b\|f(t,s)-f(\bar t,s)\|ds=0, \tag{8}$$

正数 $\lambda\in(0,1)$，那末对任给正数 ε，存在 $E_\lambda(\varepsilon)\subset[a,b]$，使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\lambda,$$
$$\lambda\int_a^b f(t,s)ds=\int_{E_\lambda(\varepsilon)}f(t,s)ds+\eta(t),$$
$$\|\eta(t)\|<\varepsilon,$$

对于 $t\in[\alpha,\beta]$ 成立.

证 因为 $[\alpha,\beta]$ 是有界闭区间，根据(8)，对于正数 ε，存在正数 $\delta>0$，使得当 $t,\bar t\in[\alpha,\beta]$，$|t-\bar t|<\delta$ 时，成立着

$$\int_a^b\|f(t,s)-f(\bar t,s)\|ds<\frac{\varepsilon}{3}. \tag{9}$$

取 $t_0=\alpha<t_1<\cdots<t_l=\beta$，使得

$$t_{j+1}-t_j<\delta \quad (j=0,\ 1,\ \cdots,\ l-1).$$

置
$$\boldsymbol{g}(s)=\begin{pmatrix}\boldsymbol{f}(t_0,\ s)\\ \boldsymbol{f}(t_1,\ s)\\ \vdots\\ \boldsymbol{f}(t_l,\ s)\end{pmatrix},$$

那末 $\boldsymbol{g}:[a,\ b]\to\mathbb{R}^{n\times(l+1)}$ 是可积的，所以存在 $E_\lambda(\varepsilon)\subset[a,\ b]$，使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\lambda,$$

$$\lambda\int_a^b \boldsymbol{g}(s)ds=\int_{E_\lambda(\varepsilon)}\boldsymbol{g}(s)ds+\boldsymbol{\eta}^*,$$

$$\|\boldsymbol{\eta}^*\|<\varepsilon/3.$$

因此
$$\lambda\int_a^b \boldsymbol{f}(t_j,\ s)ds=\int_{E_\lambda(\varepsilon)}\boldsymbol{f}(t_j,\ s)ds+\boldsymbol{\eta}_j^*,$$

$$\|\boldsymbol{\eta}_j^*\|<\varepsilon/3, \quad (j=0,\ 1,\ \cdots,\ l). \tag{10}$$

对于 $t\in[\alpha,\ \beta]$，存在 $t_j\in[\alpha,\ \beta]$ 使得

$$|t-t_j|<\delta,$$

根据(9)，成立着

$$\lambda\int_a^b \boldsymbol{f}(t,\ s)ds=\lambda\int_a^b \boldsymbol{f}(t_j,\ s)ds+\boldsymbol{\eta}_1,\quad \|\boldsymbol{\eta}_1\|<\varepsilon/3,$$

$$\int_{E_\lambda(\varepsilon)}\boldsymbol{f}(t,\ s)ds=\int_{E_\lambda(\varepsilon)}\boldsymbol{f}(t_j,\ s)ds+\boldsymbol{\eta}_2,\quad \|\boldsymbol{\eta}_2\|<\varepsilon/3.$$

由上面两式及(10)得到

$$\lambda\int_a^b \boldsymbol{f}(t,\ s)ds=\int_{E_\lambda(\varepsilon)}\boldsymbol{f}(t,\ s)ds+\boldsymbol{\eta}(t),$$

$$\|\boldsymbol{\eta}(t)\|=\|\boldsymbol{\eta}_1+\boldsymbol{\eta}_j^*-\boldsymbol{\eta}_2\|<\varepsilon. \qquad \blacksquare$$

系 设对于 $t\in[a,\ b]$，$\boldsymbol{f}(t,\ \cdot):[a,\ b]\to\mathbb{R}^n$ 是可积的，并且

$$\lim_{\bar{t}\to t}\int_a^b \|\boldsymbol{f}(t,\ s)-\boldsymbol{f}(\bar{t},\ s)\|ds=0,$$

正数 $\lambda\in(0,\ 1)$，那末对于 $\varepsilon>0$ 存在 $E_\lambda(\varepsilon)\subset[a,\ b]$，使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\lambda,$$

$$\lambda\int_a^t \boldsymbol{f}(t,\ s)ds=\int_{[a,t]\cap E_\lambda(\varepsilon)}\boldsymbol{f}(t,\ s)ds+\boldsymbol{\eta}(t),$$

$$\|\boldsymbol{\eta}(t)\|<\varepsilon. \tag{11}$$

证　置 $\boldsymbol{g}(t,\ s)=\begin{cases}\boldsymbol{f}(t,\ s), & \text{当 } a\leqslant s\leqslant t\leqslant b \text{ 时},\\ 0\quad, & \text{当 } a\leqslant t<s\leqslant b \text{ 时}.\end{cases}$

那末 $\boldsymbol{g}(t,\ s)$ 适合定理 4 的条件，从而结论得证．∎

特别，设 $\boldsymbol{f}:[a,\ b]\to\mathbb{R}^n$ 是可积的，对它应用上述系，就得到：对于 $\varepsilon>0$，存在 $E_\lambda(\varepsilon)\subset[a,\ b]$，使得

$$\mu(E_\lambda(\varepsilon))=(b-a)\lambda,$$

$$\lambda\int_a^t \boldsymbol{f}(s)ds=\int_{[a,t]\cap E_\lambda(\varepsilon)}\boldsymbol{f}(s)ds+\boldsymbol{\eta}(t),$$

$$\|\boldsymbol{\eta}(t)\|<\varepsilon.$$

对于可积函数序列 $\{\boldsymbol{f}_k\}$，成立着下述控制收敛定理．

定理 5　设 $\boldsymbol{f}_k:[a,\ b]\to\mathbb{R}^n$ 是可积的，并且存在 $m:[a,\ b]\to\mathbb{R}$ 是可积的，使得

$$\|\boldsymbol{f}_k(t)\|\leqslant m(t)\ (t\in[a,\ b],\ k=1,\ 2,\ \cdots),$$

如果等式

$$\lim_{k\to\infty}\boldsymbol{f}_k(t)=\boldsymbol{f}(t)$$

除零集外成立，那末 $\boldsymbol{f}:[a,\ b]\to\mathbb{R}^n$ 是可积的，并且

$$\int_a^b \boldsymbol{f}(t)dt=\lim_{k\to\infty}\int_a^b \boldsymbol{f}_k(t)dt.$$

该定理的证明省略了．

今后，如果某种关系除实轴上零集外是成立的，就称该关系是几乎处处成立的．例如，如果

$$\lim_{k\to\infty}f_k(t))=f(t)$$

在 $[a,\ b]$ 上除零集外是成立的，就称 f_k 在 $[a,\ b]$ 上几乎处处收敛于 f．

又，今后 $\mathbb{R}^n$ 中的元素简记为 x，不再特别记为 $\boldsymbol{x}$ 了. 从而，如果 $f:\mathbb{R}^m\to\mathbb{R}^n$，就是指 f 的定义域是 $\mathbb{R}^m$，值域是 $\mathbb{R}^n$.

定理 6 设 $f:[a, b]\times\mathbb{R}^n\to\mathbb{R}^n$ 具有性质

1° 对于 $x\in\mathbb{R}^n$，$f(\cdot, x):[a, b]\to\mathbb{R}^n$ 是可积的；

2° 对于 $t\in[a, b]$，$f(t, \cdot):\mathbb{R}^n\to\mathbb{R}^n$ 是连续的；

3° 对于 $\mathbb{R}^n$ 中的有界集 K，存在勒贝格可积的 $m:[a,b]\to\mathbb{R}$，使得不等式

$$\|f(t, x)\|\leqslant m(t)$$

对于 $t\in[a, b]$，$x\in K$ 成立.

又设 $\varphi:[a, b]\to\mathbb{R}^n$ 是连续的.

那末，$f(\cdot, \varphi(\cdot)):[a, b]\to\mathbb{R}^n$ 是可积的.

证 因为 $\varphi:[a, b]\to\mathbb{R}^n$ 是连续的，所以存在有界集 K，使得当 $t\in[a, b]$ 时

$$\varphi(t)\in K,$$

并且存在简单函数序列 $\varphi_k:[a, b]\to K$，使得

$$\|\varphi_k(t)-\varphi(t)\|<\frac{1}{k}$$

在 $t\in[a, b]$ 时成立.

设
$$\varphi_k(t)=\sum_{j=0}^{l_k-1} c_j^k\chi_{[t_j^k, t_{j+1}^k)}(t).$$

这里 $c_j^k\in\mathbb{R}^n$，那末

$$f(t, \varphi_k(t))=f\left(t, \sum_{j=0}^{l_k-1} c_j^k\chi_{[t_j^k, t_{j+1}^k)}(t)\right)$$
$$=\sum_{j=0}^{l_k-1} f(t, c_j^k)\chi_{[t_j^k, t_{j+1}^k)}(t).$$

上式右端的每一项是勒贝格可积的，所以 $f(\cdot, \varphi(\cdot))$ 是可积的，并且

$$\|f(t, \varphi_k(t))\|\leqslant m(t).$$

由 $f(t, \cdot):\mathbb{R}^n\to\mathbb{R}^n$ 是连续的，推知

$$\lim_{k\to\infty} f(t,\ \varphi_k(t)) = f(t,\ \varphi(t)).$$

根据定理 5，得证 $f(\cdot,\ \varphi(\cdot))$ 是勒贝格可积的。 ▌

习　题

1　设 φ 和 ψ 都是 $[a,\ b]$ 上的简单函数，试证明：

$$\max(\varphi,\ \psi) \text{ 和 } \min(\varphi,\ \psi)$$

都是 $[a,\ b]$ 上的简单函数。

2　设 $\{\varphi_k\}$ 是 $[a,\ b]$ 上的简单函数列，并且

$$\varphi_1(t) \geqslant \varphi_2(t) \geqslant \cdots \geqslant \varphi_k(t) \geqslant \cdots,$$

$$\varphi_k(t) \geqslant 0 \qquad (k=1,\ 2,\ \cdots)$$

如果几乎处处成立着

$$\lim_{k\to\infty} \varphi_k(t) = 0,$$

那末
$$\lim_{k\to\infty} \int_a^b \varphi_k(t)\,dt = 0.$$

3　设 $\{\varphi_k\}$ 和 $\{\psi_k\}$ 都是 $[a,\ b]$ 上的简单函数序列，并且

$$\varphi_1(t) \leqslant \varphi_2(t) \leqslant \cdots \leqslant \varphi_k(t) \leqslant \cdots,$$

$$\psi_1(t) \leqslant \psi_2(t) \leqslant \cdots \leqslant \psi_k(t) \leqslant \cdots,$$

又几乎处处成立着

$$\lim_{k\to\infty} \varphi_k(t) = \lim_{k\to\infty} \psi_k(t),$$

试证明
$$\lim_{k\to\infty} \int_a^b \varphi_k(t)\,dt = \lim_{k\to\infty} \int_a^b \psi_k(t)\,dt.$$

4　设 f、$g \in C_1$，并且几乎处处成立着

$$f(t) \leqslant g(t),$$

试证明
$$\int_a^b f(t)\,dt \leqslant \int_a^b g(t)\,dt.$$

5　设 f 在 $[a,\ b]$ 上是勒贝格可积的，则对任给正数 ε，存在连续函数 $g(\cdot)$，使得

$$\int_a^b |f(t) - g(t)|\,dt < \varepsilon.$$

6　设 f 在 $[a,\ b]$ 上是勒贝格可积的，并当 $t \in [a,\ b]$ 时规定 $f(t)$ 为零，试证

$$\lim_{h\to 0} \int_a^b |f(t+h) - f(t)|\,dt = 0.$$

7　设 f 是 $[a, b]$ 上的勒贝格可积函数，在 $[a, b]$ 外面 f 为零，试证

$$f_h(t)=\frac{1}{2h}\int_{t-h}^{t+h} f(u)du,$$

是 $[a, b]$ 上的连续函数，并且

$$\lim_{h\to 0}\int_a^b |f(t)-f_h(t)|dt=0.$$

8　设 $f:[a, b]\to\mathbb{R}^n$ 是勒贝格可积的，试证：集

$$R=\left\{x=\int_E f(t)dt \,\middle|\, E\text{ 是 }[a, b]\text{ 的可测子集}\right\}$$

的闭包是凸的(关于凸集的定义参看第二章 § 2)。

9　设 $f:[a, b]\to\mathbb{R}$ 是可积的，试证：对于 $\varepsilon>0$，存在取值为有理数、分点也为有理数的简单函数 ψ，使得

$$\int_a^b |f(t)-\psi(t)|dt<\varepsilon.$$

10　把第 9 题的结论推广到 $f:[a, b]\to\mathbb{R}^n$ 可积的情形。

11　设 $f:[a, b]\to\mathbb{R}$ 是可积的，并且

$$\int_a^b |f(t)|dt=0,$$

那末几乎处处成立

$$f(t)=0.$$

12　设 $f:[a, b]\to\mathbb{R}$ 是可积的，$f(t)\geqslant 0$，E 是可测集，则

$$\int_E f(t)dt\geqslant 0.$$

§ 2　勒贝格不定积分

设 f 是 $[a, b]$ 上的勒贝格可积函数，C 是任一常数，由

$$F(x)=\int_a^x f(t)dt+C$$

定义的函数 F 称为 f 的一个不定积分。

为讨论不定积分的性质，先引述勒贝格关于单调函数的导数定理。

定理 1　设 g 是 $[a, b]$ 上的单调增函数，即当 $a\leqslant t_1<t_2\leqslant b$ 时，$g(t_1)\leqslant g(t_2)$，且 g 是连续的，那末 g 在 $[a, b]$ 上几乎处处有导

数 $\frac{d}{dt}g(t)\geqslant 0$.

证明见夏道行等编著《实变数函数论与泛函分析概要》[3]①.

定理 2 设 g 在 $[a, b]$ 上是单调增的连续函数，那末 $\frac{dg}{dt}$ 在 $[a, b]$ 上是勒贝格可积的，并且

$$\int_a^b \frac{dg}{dt}dt \leqslant g(b)-g(a).$$

证明同定理 1.

根据上述定理，如果 g 是 $[a, b]$ 上的勒贝格可积函数，并且 $g(t)\geqslant 0$，那末

$$\int_a^t g(s)ds$$

是单调增的连续函数，并且

$$\int_a^b \frac{d}{dt}\left(\int_a^t g(s)\,ds\right)dt \leqslant \int_a^b g(t)dt. \tag{1}$$

一般的，如果 g 在 $[a, b]$ 上勒贝格可积，置

$$g^+(t)=(|g(t)|+g(t))/2,$$
$$g^-(t)=(|g(t)|-g(t))/2,$$

那末 $g^+(t)\geqslant 0$, $g^-(t)\geqslant 0$，并且 g^+、g^- 在 $[a, b]$ 上勒贝格可积，且

$$|g(t)|=g^+(t)+g^-(t),$$
$$g(t)=g^+(t)-g^-(t).$$

由是 $$\frac{d}{dt}\int_a^t g(s)\,ds=\frac{d}{dt}\int_a^t g^+(s)ds-\frac{d}{dt}\int_a^t g^-(s)ds,$$

所以 $$\int_a^b \left|\frac{d}{dt}\int_a^t g(s)ds\right|dt$$

$$\leqslant \int_a^b\left(\frac{d}{dt}\int_a^t g^+(s)\,ds+\frac{d}{dt}\int_a^t g^-(s)\,ds\right)dt,$$

① [3] 表示参考书籍序号.

从而 $$\int_a^b\left|\frac{d}{dt}\int_a^t g(s)\,ds\right|dt$$

$$\leqslant\int_a^b g^+(t)\,dt+\int_a^b g^-(t)\,dt\leqslant\int_a^b|g(t)|\,dt. \tag{2}$$

定理 3 设 f 在 $[a, b]$ 上是勒贝格可积的，那末几乎处处成立着

$$\frac{d}{dt}\int_a^t f(s)ds=f(t). \tag{3}$$

证 根据 § 1 定理 1，对于 $\varepsilon>0$，存在简单函数 φ，使得

$$\int_a^b|f(t)-\varphi(t)|dt<\varepsilon.$$

但是
$$\frac{d}{dt}\int_a^t\varphi(s)ds=\varphi(t)$$

除去有限个点外成立。因此，由(2)得到

$$\int_a^b\left|\frac{d}{dt}\int_a^t f(s)ds-f(t)\right|dt$$

$$=\int_a^b\left|\frac{d}{dt}\int_a^t(f(s)-\varphi(s))ds-(f(t)-\varphi(t))\right|dt$$

$$\leqslant\int_a^b\left|\frac{d}{dt}\int_a^t(f(s)-\varphi(s))\,ds\right|dt+\int_a^b|f(t)-\varphi(t)|\,dt$$

$$\leqslant 2\int_a^b|f(t)-\varphi(t)|\,dt<2\varepsilon.$$

根据 $\varepsilon>0$ 的任意性，由上式得

$$\int_a^b\left|\frac{d}{dt}\int_a^t f(s)ds-f(t)\right|dt=0.$$

再根据 § 1 习题 11 知(3)几乎处处成立. ▮

由等式(3)得到

$$\int_a^b\left(\frac{d}{dt}\int_a^t f(s)ds\right)dt=\int_a^b f(t)dt,$$

即

$$\int_a^b F'(t)\,dt=F(b)-F(a). \tag{4}$$

这就是不定积分的牛顿-莱布尼兹公式.

定理 4 设 $f:[a, b]\to\mathbb{R}$ 是勒贝格可积的，那末对于 $\varepsilon>0$，存在 $\delta>0$，使得当 $E\subset[a, b]$，$\mu(E)<\delta$ 时

$$\left|\int_E f(t)dt\right|<\varepsilon.$$

证 对 $\varepsilon>0$，存在简单函数 φ，使得

$$\int_a^b |f(t)-\varphi(t)|dt<\frac{\varepsilon}{2}.$$

φ 在 $[a, b]$ 上是有界的，所以存在 $\delta>0$，使得当 $\mu(E)<\delta$ 时

$$\int_E|\varphi(t)|dt<\frac{\varepsilon}{2}.$$

从而由
$$\int_E f(t)dt=\int_E(f(t)-\varphi(t))dt+\int_E\varphi(t)dt$$

得到
$$\left|\int_E f(t)dt\right|\leqslant\int_E|f(t)-\varphi(t)|dt+\int_E|\varphi(t)|dt<\varepsilon.$$

该定理表明勒贝格不定积分是绝对连续函数. ∎

定义 设 $f:[a, b]\to\mathbb{R}$ 是勒贝格可积的，$t_0\in[a, b]$，如果对于 $h_1\geqslant0$，$h_2\geqslant0$，$h=h_1+h_2>0$，成立着

$$\lim_{h\to0}\frac{1}{h}\int_{t_0-h_1}^{t_0+h_2}|f(t)-f(t_0)|dt=0, \tag{5}$$

那末称 t_0 是 f 的勒贝格点.

定理 5 设 $f:[a, b]\to\mathbb{R}$ 是可积的，t_0 是 f 的勒贝格点，那末

$$\frac{d}{dt}\int_a^t f(s)ds\bigg|_{t=t_0}=f(t_0),$$

即勒贝格点是使(1)成立的点.

证 事实上，当 $h>0$ 时，根据(5)有

$$\lim_{h\to0}\left|\frac{1}{h}\left(\int_a^{t_0+h}f(s)ds-\int_a^{t_0}f(s)ds\right)-f(t_0)\right|$$
$$=\lim_{h\to0}\left|\frac{1}{h}\int_{t_0}^{t_0+h}(f(s)-f(t_0))ds\right|$$

$$\leqslant \lim_{h\to 0} \frac{1}{h}\int_{t_0}^{t_0+h} |f(s)-f(t_0)|ds=0.$$

类似地可以考察 $h<0$ 的情况. ▌

定理 6 设 $f:[a, b]\to\mathbb{R}$ 是勒贝格可积的, 那末 $[a, b]$ 上几乎每个点是勒贝格点.

证 设 r 是有理数, 显然 $|f-r|$ 在 $[a, b]$ 上是勒贝格可积的, 记 E_r 是使(3)对 $|f-r|$ 成立的点全体, 即

$$\frac{d}{dt}\int_a^t |f(s)-r|ds=|f(t)-r|, \quad t\in E_r. \tag{6}$$

根据定理 3, $\mu(E_r)=b-a$, 即 $F_r=[a, b]\setminus E_r$ 是零集. 当 r 取遍有理数时, $\bigcup F_r=F$ 是零集, 从而 $E=\bigcap E_r$ 的测度

$$\mu(E)=b-a.$$

因此, 当 $t\in E$ 时, (6)对任何有理数 r 成立.

现证当 $t_0\in E$ 时, t_0 是 f 的勒贝格点. 不妨设 $f(t_0)$ 有限. 对 $\varepsilon>0$, 取有理数 r_0, 使

$$|f(t_0)-r_0|<\frac{\varepsilon}{2}.$$

因此, 对于 $h_1\geqslant 0$, $h_2\geqslant 0$, $h=h_1+h_2>0$, 有

$$\frac{1}{h}\int_{t_0-h_1}^{t_0+h_2} |f(t)-f(t_0)|dt$$

$$\leqslant \frac{1}{h}\int_{t_0-h_1}^{t_0+h_2} |f(t)-r_0|dt+\frac{1}{h}\int_{t_0-h_1}^{t_0+h_2} |r_0-f(t_0)|dt.$$

上式右端等于

$$\frac{1}{h}\int_{t_0-h_1}^{t_0+h_2} |f(t)-r_0|dt+|f(t_0)-r_0|$$

$$<\frac{1}{h}\int_{t_0-h_1}^{t_0+h_2} |f(t)-r_0|dt+\frac{\varepsilon}{2}.$$

当 $h\to 0$ 时, 对上式两边取上极限, 注意到(6)得

$$\overline{\lim_{h\to 0}}\frac{1}{h}\int_{t_0-h_2}^{t_0+h_2}|f(t)-f(t_0)|dt\leqslant|f(t_0)-r_0|+\frac{\varepsilon}{2}<\varepsilon.$$

再令 $\varepsilon\to 0$，就得到(5). ∎

如果 $F(\cdot):[\alpha,\beta]\to\mathbb{R}$ 是连续的，并且对于 $\varepsilon>0$，存在 $\delta>0$，使得当 $\{(\alpha_k,\ \beta_k)\}$ 是 $[\alpha,\ \beta]$ 中一族互不相交的开区间，而且总长度 $\sum\limits_k(\beta_k-\alpha_k)<\delta$ 时，不等式

$$\sum_k|F(\beta_k)-F(\alpha_k)|<\varepsilon,$$

就称 $F(\cdot)$ 是绝对连续的．下述定理表征了绝对连续函数．

定理 7 $F(\cdot):[\alpha,\beta]\to\mathbb{R}$ 为绝对连续的充要条件是：$F(\cdot)$ 为勒贝格可积函数的不定积分． ∎

证明见夏道行等的书[8].

§3 具可积右端的常微分方程组

为研究控制问题，考虑具有可积右端的常微分方程组

$$\frac{dx}{dt}=f(t,\ x).\tag{1}$$

假设它的右端 $f:[\alpha,\ \beta]\times\mathbb{R}^n\to\mathbb{R}^n$ 具有下述性质：

1° 对于 $x\in\mathbb{R}^n$，$f(\cdot,\ x):[\alpha,\ \beta]\to\mathbb{R}^n$ 是可积的；

2° 对于 $t\in[\alpha,\ \beta]$，$f(t,\cdot):\mathbb{R}^n\to\mathbb{R}^n$ 是连续的；

3° 对于 $K>0$，存在 $[\alpha,\ \beta]$ 上的可积函数

$$m_K(\cdot):[\alpha,\ \beta]\to\mathbb{R},$$

$$l_K(\cdot):[\alpha,\ \beta]\to\mathbb{R},$$

使得当 $t\in[\alpha,\ \beta]$，$\|x\|\leqslant K$，$\|y\|\leqslant K$ 时成立着

$$\|f(t,\ x)\|\leqslant m_K(t),\tag{2}$$

$$\|f(t,\ x)-f(t,\ y)\|\leqslant l_K(t)\|x-y\|.\tag{3}$$

显然，如果 $f:[\alpha,\ \beta]\times\mathbb{R}^n\to\mathbb{R}^n$ 是连续的，并且 f 关于 x 是连续可微的，那末上述条件 1°、2°、3° 成立．

又，如果对于 $x\in\mathbb{R}^n$，$f(\cdot, x):[\alpha, \beta]\to\mathbb{R}^n$ 是逐段连续的；对于 $t\in[\alpha, \beta]$，$f(t, \cdot):\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微的，那末上述条件 1°、2°、3° 成立。

满足上述条件 1°、2°、3° 的函数 f，简称为是可积的．方程(1)的 f 满足条件 1°、2°、3° 时，称为具可积右端的微分方程组．

如果 $x:(t_1, t_2)\to\mathbb{R}^n$ 是连续的，并且成立着恒等式

$$x(t)\equiv x_0+\int_{t_0}^{t} f(s, x(s))ds,\quad t\in(t_1, t_2),$$

就称 $x=x(\cdot)$ 是方程组(1)满足初值条件

$$x(t_0)=x_0 \tag{4}$$

的解．

定理 1 设 $f:[\alpha, \beta]\times\mathbb{R}^n\to\mathbb{R}^n$ 满足上述条件 1°、2°、3°，又设给定 $(t_0, x_0)\in[\alpha, \beta]\times\mathbb{R}^n$，那末方程组(1)存在唯一的满足初值条件(4)的解．

证 利用逐次逼近法来证明．

设 $M>0$，以 x_0 为中心，M 为半径的球是

$$\|x-x_0\|\leqslant M \tag{5}$$

这时
$$\|x\|\leqslant\|x_0\|+M=K.$$

设 $h>0$ 满足条件

$$[t_0-h, t_0+h]\subset[\alpha, \beta],$$

$$\int_{t_0-h}^{t_0+h} m_K(t)dt\leqslant M.$$

置
$$x^0(t)\equiv x_0,$$

显然，它位于球(5)中．逐次定义

$$x^1(t)\equiv x_0+\int_{t_0}^{t} f(s, x^0(s))ds,$$

$$x^2(t)\equiv x_0+\int_{t_0}^{t} f(s, x^1(s))ds,$$

…………………

$$x^{k+1}(t)=x_0+\int_{t_0}^{t}f(s,\ x^k(s))ds,\tag{6}$$

··················

由于 $x^0(\cdot)$ 在 $[t_0-h,\ t_0+h]$ 上连续．根据定理 1-6 (指第一节定理 6)①，知 $f(\cdot,\ x^0(\cdot))$ 在 $[t_0-h,\ t_0+h]$ 上是可积的，从而可以定义 $x^1(\cdot)$，它在 $[t_0-h,\ t_0+h]$ 上是连续的，并且当 $t\in[t_0-h,\ t_0+h]$ 时

$$\|x^1(t)-x_0\|\leqslant\left|\int_{t_0}^{t}\|f(s,\ x^0(s))\|ds\right|\leqslant\left|\int_{t_0}^{t}m_K(s)ds\right|\leqslant M.$$

因此，$x^1(t)$ 位在球(5)中，从而可以由(6)定义 $x^2(\cdot)$，并且 $x^2(t)$ 位在球(5)中．一般地，如果 $x^k(\cdot)$ 连续并且 $x^k(t)$ 位于球(5)中，那末可以用(6)来定义 $x^{k+1}(\cdot)$，并且当 $t\in[t_0-h,\ t_0+h]$ 时

$$\|x^{k+1}(t)-x_0\|\leqslant\left|\int_{t_0}^{t}m_K(s)ds\right|\leqslant M,$$

即 $x^{k+1}(t)$ 位于球(5)中．所以，我们可以用(6)逐次定义

$$x^0(\cdot),\ x^1(\cdot),\ \cdots,\ x^k(\cdot),\ \cdots.$$

并且 $x^k(\cdot)$ 在 $[t_0-h,\ t_0+h]$ 上是连续的，$x^k(t)$ 位在球(5)中．

由(6)和(3)，得到：当 $t\in[t_0-h,\ t_0+h]$ 时

$$\|x^{k+1}(t)-x^k(t)\|\leqslant\left|\int_{t_0}^{t}l_K(s)\|x^k(s)-x^{k-1}(s)\|ds\right|$$

$$(k=1,\ 2,\ \cdots)$$

由于 $$\|x^1(t)-x_0\|\leqslant M,$$

所以 $$\|x^2(t)-x^1(t)\|\leqslant M\left|\int_{t_0}^{t}l_K(s)ds\right|.$$

为确定起见，设 $t_0\leqslant t\leqslant t_0+h$，由上式得到

$$\|x^2(t)-x^1(t)\|\leqslant M\int_{t_0}^{t}l_K(s)ds.$$

如果已经证得

---

① 定理 1-6 表示本章 §1 定理 6。

$$\|x^k(t)-x^{k-1}(t)\|\leqslant\frac{M}{(k-1)!}\left(\int_{t_0}^{t}l_K(s)ds\right)^{k-1},\qquad(7)$$

那末，当 $t_0\leqslant t\leqslant t_0+h$ 时

$$\|x^{k+1}(t)-x^k(t)\|\leqslant\frac{M}{(k-1)!}\int_{t_0}^{t}l_K(s)\left(\int_{t_0}^{s}l_K(u)du\right)^{k-1}ds$$

$$=\frac{M}{k!}\left(\int_{t_0}^{t}l_K(s)ds\right)^{k}.$$

因此，对于 $t_0\leqslant t\leqslant t_0+h$，不等式(7)对于 $k=1, 2, \cdots$成立．从而

$$\|x^k(t)-x^{k-1}(t)\|\leqslant\frac{M}{(k-1)!}\left(\int_{t_0}^{t_0+h}l_K(s)ds\right)^{k-1}.$$

同样可证：当 $t\in[t_0-h, t_0]$ 时

$$\|x^k(t)-x^{k-1}(t)\|\leqslant\frac{M}{(k-1)!}\left(\int_{t_0-h}^{t_0}l_K(s)ds\right)^{k-1}.$$

所以，当 $t\in[t_0-h, t_0+h]$ 时

$$\|x^{k+1}(t)-x^k(t)\|\leqslant\frac{M}{k!}\left(\int_{t_0-h}^{t_0+h}l_K(s)ds\right)^{k}$$

$$(k=0, 1, 2, \cdots).$$

由于级数

$$\sum_{k=0}^{\infty}\frac{M}{k!}\left(\int_{t_0-h}^{t_0+h}l_K(s)ds\right)^{k}$$

是收敛的，所以函数项级数

$$x^0(t)+\sum_{k=0}^{\infty}\{x^{k+1}(t)-x^k(t)\}$$

在 $[t_0-h, t_0+h]$ 上是一致收敛的．设它的和为 $x(t)$，那末 $x(\cdot)$ 在 $[t_0-h, t_0+h]$ 上是连续的，并且在 $[t_0-h, t_0+h]$ 上一致地成立着

$$\lim_{k\to\infty}x^k(t)=x(t).$$

由于 $\|f(t, x(t))-f(t, x^k(t))\|\leqslant l_K(t)\|x(t)-x^k(t)\|$，

得到 $$\lim_{k\to\infty}f(t, x^k(t))=f(t, x(t)).$$

利用定理1-5,在(6)中令$k\to+\infty$,得到

$$x(t)=x_0+\int_{t_0}^{t}f(s,\ x(s))ds,$$

即$x(\cdot)$在$[t_0-h,\ t_0+h]$上是方程组(1)的满足初值条件(4)的解.这证明了解的存在性.

现证唯一性.

设$\bar{x}(t)$在$[t_0-h,\ t_0+h]$上是(1)的满足条件(4)的另一解.因为$\bar{x}(\cdot)$在$[t_0-h,\ t_0+h]$上是连续的,所以是有界的.从而存在$\overline{K}\geqslant K$,使得当$t\in[t_0-h,\ t_0+h]$时

$$\|\bar{x}(t)\|\leqslant\overline{K}. \tag{8}$$

还成立着

$$\|x(t)\|\leqslant\overline{K}.$$

因为

$$x(t)=x_0+\int_{t_0}^{t}f(s,\ x(s))ds,$$

$$\bar{x}(t)=x_0+\int_{t_0}^{t}f(s,\ \bar{x}(s))ds,$$

所以

$$\|x(t)-\bar{x}(t)\|\leqslant\left|\int_{t_0}^{t}l_{\overline{K}}(s)\|x(s)-\bar{x}(s)\|ds\right|. \tag{9}$$

由于$\|x(t)\|\leqslant\overline{K}$和(8),由上式得到

$$\|x(t)-\bar{x}(t)\|\leqslant2\overline{K}\left|\int_{t_0}^{t}l_{\overline{K}}(s)ds\right|.$$

再把上式代入(9)的右端,就得到

$$\|x(t)-\bar{x}(t)\|\leqslant2\overline{K}\left|\int_{t_0}^{t}l_{\overline{K}}(s)\left|\int_{t_0}^{s}l_{\overline{K}}(u)du\right|ds\right|$$

$$=\overline{K}\left(\left|\int_{t_0}^{t}l_{\overline{K}}(s)ds\right|\right)^2.$$

一般地,可以逐次证明:当$t\in[t_0-h,\ t_0+h]$时

$$\|x(t)-\bar{x}(t)\|\leqslant\frac{2\overline{K}}{k!}\left(\left|\int_{t_0}^{t}l_{\overline{K}}(s)ds\right|\right)^k.$$

$$(k=1,\ 2,\ \cdots).$$

上式右端当 $k\to\infty$ 时趋于零，所以

$$\bar{x}(t)\equiv x(t).$$ ∎

该定理是常微分方程有关定理的推广，证明方法也类似。

习　　题

1　对于线性微分方程组

$$\frac{dx}{dt}=A(t)x+f(t),\quad x(t_0)=x_0.$$

假设 $A(\cdot):[\alpha,\beta]\to\mathbb{R}^{n\times n}$, $f(\cdot):[\alpha,\beta]\to\mathbb{R}^n$ 是可积的，试证：它的解在 $[\alpha,\beta]$ 上存在且唯一，这里 $t_0\in[\alpha,\beta]$.

2　对于微分方程组

$$\frac{dx}{dt}=f(t,x), \tag{1}$$

假设 $f:[\alpha,\beta]\times\mathbb{R}^n\to\mathbb{R}^n$ 具有性质 1°、2° 及

3　对于 $K>0$，存在 $[\alpha,\beta]$ 上可积的 $m_K(\cdot)$，使得

$$\|f(t,x)\|\leqslant m_K(t)$$

对于 $t\in[\alpha,\beta]$, $\|x\|\leqslant K$ 成立．又设 $(t_0,x_0)\in[\alpha,\beta]\times\mathbb{R}^n$ 给定。

试证方程组(1)存在满足初值条件(4)的解.

§4　解对参数的连续依赖性

考虑右端依赖于参数 μ 的微分方程组

$$\frac{dx}{dt}=f(t,x,\mu), \tag{1}$$

假设它的右端 $f:[\alpha,\beta]\times\mathbb{R}^n\times(0,1)\to\mathbb{R}^n$ 满足以下条件：

1°　对于 $(x,\mu)\in\mathbb{R}^n\times(0,1)$, $f(\cdot,x,\mu):[\alpha,\beta]\to\mathbb{R}^n$ 是可积的；

2°　对于 $(t,\mu)\in[\alpha,\beta]\times(0,1)$, $f(t,\cdot,\mu):\mathbb{R}^n\to\mathbb{R}^n$ 是连续的；

3°　对于 $K>0$，存在 $[\alpha,\beta]$ 上的可积函数 $m_K(\cdot)$ 和 $l_K(\cdot)$，使得：当 $t\in[\alpha,\beta]$, $\mu\in(0,1)$, $\|x\|\leqslant K$, $\|y\|\leqslant K$ 时，成立着不等式

$$\|f(t, x, \mu)\| \leqslant m_K(t),$$
$$\|f(t, x, \mu)-f(t, y, \mu)\| \leqslant l_K(t)\|x-y\|;$$

4° 存在 $F(t, x):[\alpha, \beta]\times\mathbb{R}^n\to\mathbb{R}^n$ 满足条件：

i) 当 $x\in\mathbb{R}^n$ 时，$F(\cdot, x):[\alpha, \beta]\to\mathbb{R}^n$ 是可积的；

ii) 当 $t\in[\alpha, \beta]$ 时，$F(t, \cdot):\mathbb{R}^n\to\mathbb{R}^n$ 是连续的；

iii) 当 $t\in[\alpha, \beta]$，$\|x\|\leqslant K$ 时

$$\|F(t, x)\|\leqslant m_K(t);$$

iv) 对于 $(t, x)\in[\alpha, \beta]\times\mathbb{R}^n$，成立着

$$\lim_{\mu\to 0^+}\int_{t_0}^{t} f(s, x, \mu)ds=\int_{t_0}^{t} F(s, x)ds, \tag{2}$$

这里 $t_0\in[\alpha, \beta]$ 给定；

5° 微分方程组

$$\frac{dx}{dt}=F(t, x) \tag{3}$$

满足初值条件

$$x(t_0)=x_0 \tag{4}$$

的解 $x=\xi(t)$ 在 $[t_0, t_0+h]$ 上存在且唯一，这里 $x_0\in\mathbb{R}^n$ 给定，$h>0$ 可以类似于 § 3 来确定.

本节讨论方程组(1)满足初值条件(4)的解

$$x=x(t, \mu)$$

与方程组(3)满足初值条件(4)的解之间的关系.

我们假设 f 和 F 满足上述条件 1°—5°. 现在，先证明下述的引理.

引理 1 设 $x(\cdot)$ 是 $[\alpha, \beta]$ 上的简单函数，那末对于 $t\in[\alpha, \beta]$，等式

$$\lim_{\mu\to 0^+}\int_{t_0}^{t} f(s, x(s), \mu)ds=\int_{t_0}^{t} F(s, x(s))ds \tag{5}$$

成立. 并且，对于 $\varepsilon>0$，存在 $\delta>0$，使得当 $0<\mu<\delta$ 时，不等式

$$\left\|\int_{t_0}^{t} f(s,\ x(s),\ \mu)ds-\int_{t_0}^{t} F(s,\ x(s))ds\right\|<\varepsilon \tag{6}$$

对于 $t\in[\alpha,\ \beta]$ 成立.

证　等式(5)是等式(2)的直接推论.

其次，对于 $t,\ t'\in[\alpha,\ \beta]$，根据条件 3°，成立着

$$\left\|\int_{t}^{t'} f(s,\ x(s),\ \mu)ds\right\|\leqslant\left|\int_{t}^{t'} m_K(s)ds\right|,$$

这里 $K=\max\limits_{\alpha\leqslant t\leqslant\beta}\|x(t)\|$.

在上式中令 $\mu\to0^+$，并根据(5)得到

$$\left\|\int_{t}^{t'} F(s,\ x(s))ds\right\|\leqslant\left|\int_{t}^{t'} m_K(s)ds\right|.$$

因为 $m_K(\cdot)$ 在 $[\alpha,\ \beta]$ 上勒贝格可积，所以存在正数 $\eta>0$，使得当 $|t-t'|<\eta$ 时

$$\left|\int_{t}^{t'} m_K(s)ds\right|<\frac{\varepsilon}{3},$$

(参见定理 2-4)从而

$$\left\|\int_{t}^{t'} f(s,\ x(s),\ \mu)ds\right\|<\frac{\varepsilon}{3}, \tag{7}$$

$$\left\|\int_{t}^{t'} F(s,\ x(s))ds\right\|<\frac{\varepsilon}{3}. \tag{8}$$

把区间$[\alpha,\ \beta]$用分点 $\tau_0=\alpha<\tau_1<\cdots<\tau_l=\beta$ 分割，使得

$$\tau_{j+1}-\tau_j<\eta\quad(j=0,\ 1,\ \cdots,\ l-1).$$

由于等式(5)对于 $t=\tau_0,\ \tau_1,\ \cdots,\ \tau_l$ 成立，所以存在正数 δ，使得当 $0<\mu<\delta$ 时

$$\left\|\int_{t_0}^{\tau_j} f(s,\ x(s),\ \mu)ds-\int_{t_0}^{\tau_j} F(s,\ x(s))ds\right\|<\frac{\varepsilon}{3}$$
$$(j=0,\ 1,\ 2,\ \cdots,\ l) \tag{9}$$

对于 $t\in[\alpha,\ \beta]$，存在 τ_j，使得

$$|t-\tau_j|<\eta.$$

因此，当 $0<\mu<\delta$ 时，根据(7)，(8)和(9)得到

$$\left\|\int_{t_0}^{t} f(s, x(s), \mu)ds-\int_{t_0}^{t} F(s, x(s))ds\right\|$$

$$\leqslant\left\|\int_{t_0}^{\tau_j} f(s, x(s), \mu)ds-\int_{t_0}^{\tau_j} F(s, x(s))ds\right\|$$

$$+\left\|\int_{\tau_j}^{t} f(s, x(s), \mu)ds\right\|+\left\|\int_{t}^{\tau_j} F(s, x(s))ds\right\|<\varepsilon. \quad \blacksquare$$

引理 2 设 $\zeta(\cdot):[\alpha, \beta]\to\mathbb{R}^n$ 和 $\varphi(\cdot, \mu):[\alpha, \beta]\to\mathbb{R}^n$ 是连续的，并且当 $\mu\to0^+$ 时 $\varphi(t, \mu)$ 一致收敛于 $\zeta(t)$，即

$$\lim_{\mu\to0^+}\max_{\alpha\leqslant t\leqslant\beta}\|\varphi(t, \mu)-\zeta(t)\|=0,$$

那末等式

$$\lim_{\mu\to0^+}\int_{t_0}^{t} f(s, \varphi(s, \mu), \mu)ds=\int_{t_0}^{t} F(s, \zeta(s))ds \tag{10}$$

对于 $t\in[\alpha, \beta]$ 成立，并且极限关于 t 是一致的。

证 因 $\zeta(\cdot)$ 在 $[\alpha, \beta]$ 上连续，所以

$$\max_{\alpha\leqslant t\leqslant\beta}\|\zeta(t)-x_0\|=M/2$$

存在. 置

$$K=\|x_0\|+M,$$

那末，当 $t\in[\alpha, \beta]$ 时

$$\|\zeta(t)\|\leqslant K.$$

对于 $\varepsilon>0$，存在 $\eta>0$ $(\eta<M/2)$，使得

$$\eta\int_{\alpha}^{\beta} l_K(s)ds<\frac{\varepsilon}{6}.$$

根据假设，存在 $\delta>0$，使得当 $0<\mu<\delta$ 时

$$\|\varphi(t, \mu)-\zeta(t)\|<\eta$$

对于 $t\in[\alpha, \beta]$ 成立. 从而

$$\|\varphi(t, \mu)\|\leqslant\eta+\|\zeta(t)\|\leqslant\eta+\|x_0\|+M/2<K.$$

对于 $\zeta(\cdot)$ 存在简单函数 $\xi_1(\cdot):[\alpha, \beta]\to\mathbb{R}^n$ 满足不等式

$$\max_{\alpha\leqslant t\leqslant\beta}\|\xi_1(t)-\zeta(t)\|<\eta,$$

$$\max_{\alpha\leqslant t\leqslant\beta}\|\xi_1(t)\|\leqslant K.$$

根据条件 3°，当 $x(\cdot)$、$y(\cdot)$ 是简单函数，且

$$\|x(t)\|\leqslant K,\quad \|y(t)\|\leqslant K$$

时
$$\left\|\int_{t_0}^t f(s,\ x(s),\ \mu)ds-\int_{t_0}^t f(s,\ y(s),\ \mu)ds\right\|$$
$$\leqslant\left|\int_{t_0}^t l_K(s)\|x(s)-y(s)\|ds\right|$$
$$\leqslant\max_{\alpha\leqslant\tau\leqslant\beta}\|x(\tau)-y(\tau)\|\left|\int_{t_0}^t l_K(s)ds\right|.$$

根据引理 1，令 $\mu\to0^+$，由上式得到

$$\left\|\int_{t_0}^t F(s,\ x(s))ds-\int_{t_0}^t F(s,\ y(s))ds\right\|$$
$$\leqslant\max_{\alpha\leqslant\tau\leqslant\beta}\|x(\tau)-y(\tau)\|\cdot\left|\int_{t_0}^t l_K(s)ds\right|.$$

又因为连续函数可以用简单函数一致逼近，所以上式对于 $x(\cdot)$ 或 $y(\cdot)$ 为连续函数也成立. 特别

$$\left\|\int_{t_0}^t F(s,\ \xi_1(s))ds-\int_{t_0}^t F(s,\ \zeta(s))ds\right\|$$
$$\leqslant\max_{\alpha\leqslant\tau\leqslant\beta}\|\xi_1(\tau)-\zeta(\tau)\|\cdot\left|\int_{t_0}^t l_K(s)ds\right|$$
$$<\eta\left|\int_{t_0}^t l_K(s)ds\right|<\frac{\varepsilon}{6}. \tag{11}$$

再根据 3°，我们有

$$\left\|\int_{t_0}^t f(s,\ \varphi(s,\ \mu),\ \mu)ds-\int_{t_0}^t f(s,\ \xi_1(s),\ \mu)ds\right\|$$
$$\leqslant\left|\int_{t_0}^t l_K(s)\|\varphi(s,\ \mu)-\xi_1(s)\|ds\right|$$
$$\leqslant2\eta\left|\int_{t_0}^t l_K(s)ds\right|<\frac{\varepsilon}{3}. \tag{12}$$

对于简单函数 $\xi_1(\cdot)$，应用引理 1，存在 $\delta_0>0$，使得当 $0<\mu<\delta_0(\leqslant\delta)$ 时

$$\left\|\int_{t_0}^{t} f(s,\ \xi_1(s),\ \mu)ds - \int_{t_0}^{t} F(s,\ \xi_1(s))ds\right\| < \frac{\varepsilon}{2}$$

对于 $t\in[\alpha,\ \beta]$ 成立.

从而，由上式及(11)、(12)推知，当 $0<\mu<\delta_0$ 时

$$\left\|\int_{t_0}^{t} f(s,\ \varphi(s,\ \mu),\ \mu)ds - \int_{t_0}^{t} F(s,\ \zeta(s))ds\right\|$$

$$\leqslant \left\|\int_{t_0}^{t} f(s,\ \varphi(s,\ \mu),\ \mu)ds - \int_{t_0}^{t} f(s,\ \xi_1(s),\ \mu)ds\right\|$$

$$+\left\|\int_{t_0}^{t} f(s,\ \xi_1(s),\ \mu)ds - \int_{t_0}^{t} F(s,\ \xi_1(s))ds\right\|$$

$$+\left\|\int_{t_0}^{t} F(s,\ \xi_1(s))ds - \int_{t_0}^{t} F(s,\ \xi(s))ds\right\| < \varepsilon.$$ ∎

现在证明

定理 1 设条件 1°—5° 成立，那末对于 $\varepsilon>0$ 存在 $\delta>0$，使得：当 $0<\mu<\delta$ 时，方程组(1)以 $(t_0,\ x_0)$ 为初值的解 $x=x(t,\ \mu)$ 在 $[t_0,\ t_0+h]$ 上存在，并且

$$\|x(t,\ \mu)-\xi(t)\|<\varepsilon.$$

证 设 $M>0$ 给定，$K=\|x_0\|+M$，取 $h>0$ 使得

$$\int_{t_0}^{t_0+h} m_K(s)ds\leqslant M.$$

根据定理 3-1 的证明，$x(\cdot,\ \mu)$ 在 $[t_0,\ t_0+h]$ 上存在，并且

$$\|x(t,\ \mu)-x_0\|\leqslant M,$$

从而
$$\|x(t,\ \mu)\|\leqslant K,$$

即 $\{x(\cdot,\ \mu)\}$ 是一致有界的。

其次，由

$$x(t,\ \mu)\equiv x_0+\int_{t_0}^{t} f(s,\ x(s,\ \mu),\ \mu)ds \tag{13}$$

容易得到 $\|x(t,\ \mu)-x(t',\ \mu)\|\leqslant\left|\int_{t}^{t'} m_K(s)ds\right|.$

所以 $\{x(\cdot,\ \mu)\}$ 是同等连续的.

我们要证明: $x(\cdot,\ \mu)$ 一致收敛于 $\xi(\cdot)$.

否则，存在 $r>0$ 和 $\mu_k\to 0$, 使得

$$\max_{t_0\leqslant t\leqslant t_0+h}\|x(t,\ \mu_k)-\xi(t)\|\geqslant r. \tag{14}$$

因为 $\{x(\cdot,\ \mu_k)\}$ 是一致有界且同等连续的，根据阿尔采拉(C. Arzelà)定理，存在一致收敛的子列，仍记它为$\{x(\cdot,\ \mu_k)\}$，设它的极限为$\zeta(\cdot)$.

根据引理 2, 成立着

$$\lim_{k\to\infty}\int_{t_0}^{t}f(s,\ x(s,\ \mu_k),\ \mu_k)ds=\int_{t_0}^{t}F(s,\ \zeta(s))ds.$$

在(13)中设 $\mu=\mu_k$, 并令 $k\to+\infty$, 得到

$$\zeta(t)\equiv x_0+\int_{t_0}^{t}F(s,\ \zeta(s))ds.$$

即 $x=\zeta(\cdot)$ 是方程组(3)的解，以$(t_0,\ x_0)$为初值. 根据定理的假设，方程组(3)以$(t_0,\ x_0)$为初值的解唯一，所以

$$\zeta(t)\equiv\xi(t). \tag{15}$$

但在(14)中令 $k\to\infty$, 得到

$$\max_{t_0\leqslant t\leqslant t_0+h}\|\zeta(t)-\xi(t)\|\geqslant r>0.$$

它与(15)矛盾.

因此，当 $\mu\to 0^+$ 时 $x(\cdot,\ \mu)$在$[t_0,\ t_0+h]$上一致收敛于$\xi(\cdot)$.▊

习　　题

1　设 $g:[\alpha,\ \beta]\to\mathbb{R}^n$ 是连续的，证明: 对于 $\varepsilon>0$ 存在简单函数 $\varphi:[\alpha,\ \beta]\to\mathbb{R}^n$, 满足

$$\|\varphi(t)-g(t)\|<\varepsilon,$$

$$\max_{\alpha\leqslant t\leqslant\beta}\|\varphi(t)\|\leqslant\max_{\alpha\leqslant t\leqslant\beta}\|g(t)\|.$$

2　设 $f:[\alpha,\ \beta]\times\mathbb{R}^n\times[0,\ 1)\to\mathbb{R}^n$ 连续，并且对于 $(t,\ \mu)\in[\alpha,\ \beta]\times[0,\ 1)$, $f(t,\ \cdot,\ \mu):\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微的，试验证: 条件 1°～5° 成立.

3　试讨论微分方程

$$\frac{dx}{dt}=f(t,\ x)\sin kt$$

以 (0, 0) 为初值的解 $x=\varphi_k(\cdot)$ 当 $k\to+\infty$ 时的极限是否存在? 这里 f: $[\alpha,\ \beta]\times\mathbb{R}\to\mathbb{R}$ 是连续的, 并且对于 $t\in[\alpha,\ \beta]$, $f(t,\ \cdot):\mathbb{R}\to\mathbb{R}$ 是连续可微的.

§5 线性微分方程组和它的共轭方程组

设 $A(\cdot):[\alpha,\ \beta]\to\mathbb{R}^{n\times n}$, $f:[\alpha,\ \beta]\to\mathbb{R}^n$ 是可积的, 考虑线性微分方程组

$$\frac{dx}{dt}=A(t)x+f(t), \tag{1}$$

对于 $(t_0,\ x_0)\in[\alpha,\ \beta]\times\mathbb{R}^n$, 在 $[\alpha,\ \beta]$ 上考虑逐次逼近序列

$$x^0(t)\equiv x_0,$$

$$x^1(t)\equiv x_0+\int_{t_0}^{t}\{A(s)x^0(s)+f(s)\}ds,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x^{k+1}(t)\equiv x_0+\int_{t_0}^{t}\{A(s)x^k(s)+f(s)\}ds,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

这时, $x^k(\cdot):[\alpha,\ \beta]\to\mathbb{R}^n$ 是连续的, 并且

$$\|x^1(t)-x^0(t)\|\leqslant\left|\int_{t_0}^{t}\{\|A(s)\|\,\|x^0(s)\|+\|f(s)\|\}ds\right|,$$

$$\|x^{k+1}(t)-x^k(t)\|\leqslant\left|\int_{t_0}^{t}\|A(s)\|\cdot\|x^k(s)-x^{k-1}(s)\|ds\right|.$$

可以证明: $\{x^k(\cdot)\}$ 在 $[\alpha,\ \beta]$ 上一致收敛于连续的 $x(\cdot)$, 并且

$$x(t)\equiv x_0+\int_{t_0}^{t}\{A(s)x(s)+f(s)\}ds,$$

即 (1) 以 $(t_0,\ x_0)$ 为初值的解在 $[\alpha,\ \beta]$ 上存在. 和 §3 一样, 可以证明它是唯一的.

当 $f(t)\equiv0$ 时, (1) 成为齐次线性微分方程组

$$\frac{dx}{dt}=A(t)x. \tag{2}$$

设它以$(t_0,\ x_0)$为初值的解记为$x=x(\cdot,\ t_0,\ x_0)$.

显然，当$t,\ t_0\in[\alpha,\ \beta]$时，$x(t,\ t_0,\ \cdot):\mathbb{R}^n\to\mathbb{R}^n$是线性的，即

$$x(t;\ t_0,\ \alpha_1x_0^1+\alpha_2x_0^2)=\sum_{j=1}^{2}\alpha_jx(t,\ t_0,\ x_0^j).$$

引理 1 [格隆瓦尔 (T. H. Gronwall) 不等式] 设$u(\cdot)$, $v(\cdot)$是$[t_0,\ \infty)$上的实连续函数，$v(t)\geqslant0$，如果当$t\geqslant t_0$时成立着

$$u(t)\leqslant\alpha+\int_{t_0}^{t}v(s)u(s)ds,$$

那末不等式
$$u(t)\leqslant\alpha\exp\left(\int_{t_0}^{t}v(s)ds\right)$$

对于$t\geqslant t_0$成立.

证 置
$$w(t)=\alpha+\int_{t_0}^{t}v(s)u(s)ds,$$

那末当$t\geqslant t_0$时

$$u(t)\leqslant w(t),$$

$$\frac{dw(t)}{d}=v(t)u(t)\leqslant v(t)w(t),$$

$$w(t_0)=\alpha,$$

从而
$$\frac{d}{dt}\left\{\exp\left(-\int_{t_0}^{t}v(s)ds\right)w(t)\right\}\leqslant0,$$

所以
$$\exp\left(-\int_{t_0}^{t}v(s)ds\right)w(t)\leqslant w(t_0)=\alpha,$$

因此
$$w(t)\leqslant\alpha\exp\left(\int_{t_0}^{t}v(s)ds\right).$$

从而得证
$$u(t)\leqslant w(t)\leqslant\alpha\exp\left(\int_{t_0}^{t}v(s)ds\right).$$ ∎

由于
$$x(t,\ t_0,\ x_0)=x_0+\int_{t_0}^{t}A(s)x(s,\ t_0,\ x_0)ds,$$

所以，当$t_0\leqslant t$时

$$\|x(t, t_0, x_0)\| \leq \|x_0\| + \int_{t_0}^{t} \|A(s)\| \|x(s, t_0, x_0)\| ds.$$

根据格隆瓦尔不等式，由上式得到

$$\|x(t, t_0, x_0)\| \leq \|x_0\| \exp\left(\int_{t_0}^{t} \|A(s)\| ds\right) \tag{3}$$

对于 $t \geq t_0$ 成立，从而

$$\|x(t, t_0, x_0)\| \leq \|x_0\| \exp\left(\int_{\alpha}^{\beta} \|A(s)\| ds\right) \tag{4}$$

对于 $\alpha \leq t \leq t_0$，我们可以考虑格隆瓦尔不等式的形式，从而得到：当 $\alpha \leq t \leq t_0$ 时

$$\|x(t, t_0, x_0)\| \leq \|x_0\| \exp\left(\int_{t}^{t_0} \|A(s)\| ds\right).$$

从而不等式(4)对 $t \in [\alpha, \beta]$ 都成立．这就是说，当固定 t_0, t 时，$x(t, t_0, \cdot): \mathbb{R}^n \to \mathbb{R}^n$ 是线性有界算子，它可表示为

$$x(t, t_0, x_0) \equiv \Phi(t, t_0) x_0.$$

从而
$$\Phi(t, t_0) x_0 \equiv x_0 + \int_{t_0}^{t} A(s) \Phi(s, t_0) x_0 ds$$

对于 $x_0 \in \mathbb{R}^n$ 皆成立，所以

$$\Phi(t, t_0) \equiv I + \int_{t_0}^{t} A(s) \Phi(s, t_0) ds.$$

这里 I 是单位阵，即有

$$\Phi(t_0, t_0) = I,$$

$$\frac{d}{dt} \Phi(t, t_0) \equiv A(t) \Phi(t, t_0) \tag{5}$$

我们称 $\Phi(t, t_0)$ 为方程组(2)的转移矩阵，也称为(2)的基本解方阵．

根据初值问题解的唯一性，我们有

$$x(t, s, x(s, t_0, x_0)) \equiv x(t, t_0, x_0),$$

即
$$\Phi(t, s)\{\Phi(s, t_0) x_0\} \equiv \Phi(t, t_0) x_0,$$

因此

$$\Phi(t, s)\Phi(s, t_0)\equiv\Phi(t, t_0). \tag{6}$$

特别，由上式得到

$$\Phi(t_0, s)\Phi(s, t_0)\equiv\Phi(t_0, t_0)\equiv I,$$

所以 $\Phi(s, t_0)$ 是满秩的，它的逆是 $\Phi(t_0, s)$，即

$$\Phi^{-1}(s, t_0)\equiv\Phi(t_0, s). \tag{7}$$

现考虑非齐次方程组(1)，置

$$x=\Phi(t, t_0)y, \tag{8}$$

那末 $$\Phi(t, t_0)\frac{dy}{dt}=f(t),$$

从而 $$\frac{dy}{dt}=\Phi(t_0, t)f(t),$$

因此 $$y(t)\equiv y(t_0)+\int_{t_0}^{t}\Phi(t_0, s)f(s)ds.$$

由上式和(8)，有

$$\begin{aligned} x(t)&\equiv\Phi(t, t_0)y(t_0)+\int_{t_0}^{t}\Phi(t, t_0)\Phi(t_0, s)f(s)ds \\ &\equiv\Phi(t, t_0)x_0+\int_{t_0}^{t}\Phi(t, s)f(s)ds. \end{aligned} \tag{9}$$

它是方程组(1)以 (t_0, x_0) 为初值的解，称为方程组(1)的常数变易公式. 由此看出称 Φ 为转移矩阵的理由.

对于常系数线性微分方程组

$$\frac{dx}{dt}=Ax+f(t) \tag{10}$$

和

$$\frac{dx}{dt}=Ax. \tag{11}$$

这时 $x(\cdot, t_0, x_0)$ 和 $x(\cdot-t_0, 0, x_0)$ 一致，所以

$$\Phi(t, t_0)\equiv\Phi(t-t_0, 0).$$

通常把它记为 $e^{A(t-t_0)}$，从而，对于(10)，常数变易公式为

$$x=e^{A(t-t_0)}x_0+\int_{t_0}^{t}e^{A(t-s)}f(s)ds.$$

方程组(1)的共轭方程组是下述方程组

$$\frac{d\psi}{dt}=-A^T(t)\psi+q(t). \tag{12}$$

显然,(12)的共轭方程组是(1).

设 $x(\cdot)$ 和 $\psi(\cdot)$ 分别为(1)和(12)的解,那末

$$\frac{d}{dt}\langle\psi(t),\ x(t)\rangle\equiv\langle-A^T(t)\psi(t)+q(t),x(t)\rangle$$
$$+\langle\psi(t),\ A(t)x(t)+f(t)\rangle,$$

从而 $$\frac{d}{dt}\langle\psi(t),\ x(t)\rangle\equiv\langle q(t),x(t)\rangle+\langle\psi(t),\ f(t)\rangle.$$

因此,成立着

$$\langle\psi(t_1),\ x(t_1)\rangle-\langle\psi(t_0),\ x(t_0)\rangle$$
$$=\int_{t_0}^{t_1}\{\langle q(t),\ x(t)\rangle+\langle\psi(t),\ f(t)\rangle\}dt. \tag{13}$$

(13)是联系互为共轭的方程组(1)和(12)之间关系的等式.

如果 $\Psi(t,\ t_0)$ 是(12)的齐次方程组

$$\frac{d\psi}{dt}=-A^T(t)\psi \tag{14}$$

的转移矩阵,那末 $\Psi(t,\ t_0)\psi_0$ 和 $\Phi(t,\ t_0)x_0$ 分别是 (14) 和 (2) 的解,这时

$$\frac{d}{dt}\langle\Psi(t,\ t_0)\psi_0,\ \Phi(t,\ t_0)x_0\rangle\equiv0,$$

所以 $$\langle\Psi(t,\ t_0)\psi_0,\ \Phi(t,\ t_0)x_0\rangle\equiv\langle\psi_0,\ x_0\rangle,$$

即 $$\langle\Phi^T(t,\ t_0)\Psi(t,\ t_0)\psi_0,\ x_0\rangle\equiv\langle\psi_0,\ x_0\rangle.$$

由于上式对任意 ψ_0, x_0 都成立,从而

$$\Phi^T(t,\ t_0)\Psi(t,\ t_0)\equiv I,$$

即

$$\Psi(t,\ t_0)\equiv(\Phi^T(t,\ t_0))^{-1}\equiv\Phi^T(t_0,\ t). \tag{15}$$

从而(12)的解可表示为

$$\psi=\Phi^T(t_0,t)\psi_0+\int_{t_0}^{t}\Phi^T(s,\ t)q(s)ds.$$

习　　题

1　设$f:[\alpha,\ \beta]\times\mathbb{R}^n\to\mathbb{R}^n$满足条件:

1° 对于$x\in\mathbb{R}^n$, $f(\cdot,\ x):[\alpha,\ \beta]\to\mathbb{R}^n$是可积的;

2° 对于$t\in[\alpha,\ \beta]$, $f(t,\ \cdot):\mathbb{R}^n\to\mathbb{R}^n$是连续的;

3° 存在可积的$l(\cdot):[\alpha,\ \beta]\to\mathbb{R}$, 使得: 当$t\in[\alpha,\ \beta]$, $x,\ y\in\mathbb{R}^n$时

$$\|f(t,\ x)-f(t,\ y)\|\leqslant l(t)\|x-y\|.$$

试证明: 方程组

$$\frac{dx}{dt}=f(t,\ x)$$

的以$(t_0,\ x_0)\in[\alpha,\ \beta]\times\mathbb{R}^n$为初值的解在$[\alpha,\ \beta]$上存在且唯一。

2　设$u(\cdot)$、$v(\cdot)$是$(-\infty,\ t_0]$上的实函数, $v(t)\geqslant0$, 并且成立着不等式

$$u(t)\leqslant\alpha+\int_{t}^{t_0}v(s)u(s)ds,\qquad t\leqslant t_0,$$

那末成立着不等式

$$u(t)\leqslant\alpha\exp\left(\int_{t}^{t_0}v(s)ds\right),\qquad t\leqslant t_0.$$

3　设$u(\cdot)$、$v(\cdot)$在$[t_0,\ \infty)$上是连续可微的, $p(\cdot)$和$q(\cdot)$是$[t_0,\ \infty)$上的连续函数, 成立着

$$\frac{dv(t)}{dt}=p(t)v(t)+q(t),$$

$$\frac{du(t)}{dt}\leqslant p(t)u(t)+q(t),$$

$$v(t_0)=u(t_0),$$

试证明: 成立着不等式

$$u(t)\leqslant v(t).$$

第二章　线性系统的最优控制

§1　引　　言

线性系统

$$\frac{dx}{dt}=Ax+Bu(t),\quad u(t)\in U,$$

$$x(t_0)=x_0,\quad x(t_1)=x_1$$

的时间最优控制问题，是在 A、B、U、x_0、x_1 给定条件下，选择容许控制，使得上述条件成立，并且时间 t_1-t_0 达到最小.

这是最优控制系统设计中最早讨论的问题. 契贝晓夫是第一个讨论该问题的学者. 但是，把这类问题系统地做为控制问题进行讨论，正如绪论中所提到的，第一个工作是布绍 1953 年关于二阶线性系统最优开关控制的研究，这是在数学家列甫希茨的推动下进行的. 钱学森在 1954 年出版的《工程控制论》[1]中引述了布绍的工作.

其后，苏联的自动控制理论学家费里德堡 (A. A. Фельдбаум) 得到线性系统的最优开关次数定理，庞特里雅金等利用凸集理论证明了时间最优控制的最大值原理，克拉索夫斯基利用矩量问题方法讨论该问题. 在美国，拉萨尔利用向量值测度证明开关定理，值尔曼等讨论了最优开关控制.

在我国，宋健等利用等时区域方法给出了时间最优控制的综合方法，李训经等利用向量值测度证明：等时区域在 U 为任一集时也是凸的，并给出一般情形下最优时间的计算公式.

应当说，在本世纪五、六十年代，线性系统的时间最优控制是非常活跃的领域，吸引了许多数学家和控制理论学家. 由于这是

一个不能用古典变分学研究的问题，它的研究和解决为一般形式的最大值原理提供了基础，并且它也是最大值原理的应用实例，所以线性系统的时间最优控制成为最优控制理论的出发点和有力支柱.

线性系统最优反馈调节器的设计问题，是一个可以用二次泛函的极值问题来解决的问题，在线性控制系统理论课中一般都要讲述. 本章也将简述其主要结果.

本章首先讲述有限维线性空间中的凸集，特别是关于凸集的分离性定理，接着讨论等时区域的定义和它的凸闭性，利用等时区域的凸闭性推导最大值原理、最优开关原理和最优时间的计算公式. 最后，讨论线性系统的一般形式的最优控制问题，它将成为以后证明一般形式的最大值原理的基础.

§2 凸 集

对于 $\mathbb{R}^n$ 中的点 $x=(x^1, x^2, \cdots, x^n)$ 和 $y=(y^1, y^2, \cdots, y^n)$，我们规定它们的内积为

$$\langle x, y\rangle=\sum_{j=1}^{n} x^j y^j.$$

点 x 的范数为

$$\|x\|=\{\langle x, x\rangle\}^{\frac{1}{2}}=\left\{\sum_{j=1}^{n}(x^j)^2\right\}^{\frac{1}{2}}.$$

定义 1 设 Ω 是 $\mathbb{R}^n$ 中的点集. 如果对于 $x_j\in\Omega(j=1, 2)$ 和 $\alpha\in[0, 1]$，点 $\alpha x_1+(1-\alpha)x_2\in\Omega$，即以 x_1 和 x_2 为端点的线段完全在 Ω 中，就称 Ω 是凸集.

显然，若干个凸集的交集是凸集.

两个凸集 Ω_1 和 Ω_2 的差集 $\Omega=\Omega_1-\Omega_2=\{x=x_1-x_2 \mid x_1\in\Omega_1, x_2\in\Omega_2\}$ 是凸集.

两个凸集的和集 $\Omega_1+\Omega_2$ 也是凸集.

但两个凸集的并集 $\Omega_1\cup\Omega_2$ 可能是不凸的.

定理1 设 Ω_1 和 Ω_2 是 $\mathbb{R}^n$ 中的凸闭集，Ω_1 或 Ω_2 是有界的，那末 $\Omega=\Omega_1-\Omega_2$ 是凸闭集.

证 已证 Ω 的凸性，现证它的闭性.

设 $x_0\in\bar{\Omega}$，即存在 $x_k\in\Omega$，使得

$$x_0=\lim_{k\to\infty}x_k. \tag{1}$$

因为 $x_k\in\Omega$，所以存在 $x_k^{(j)}\in\Omega_j(j=1,\ 2)$，使得

$$x_k=x_k^{(1)}-x_k^{(2)}. \tag{2}$$

设 Ω_1 是有界的，从而 $\{x_k^{(1)}\}$ 是有界的，它有收敛子列，不妨设是它本身，它的极限记为 $x^{(1)}$. 因 Ω_1 是闭的，所以

$$x^{(1)}\in\Omega_1.$$

由(1)，(2)和$\{x_k^{(1)}\}$的极限为 $x^{(1)}$，得到

$$\lim_{k\to\infty}x_k^{(2)}=\lim_{k\to\infty}x_k^{(1)}-\lim_{k\to\infty}x_k=x^{(1)}-x_0.$$

从而 $x^{(2)}=x^{(1)}-x_0\in\Omega_2$（因 Ω_2 是闭的）. 所以

$$x_0=x^{(1)}-x^{(2)}\in\Omega. \qquad \blacksquare$$

定理2 设 Ω 是 $\mathbb{R}^n$ 中的凸闭集，$\bar{x}$ 是 $\mathbb{R}^n$ 中给定的点，那末存在唯一的 $x_0\in\Omega$，使得

$$\|\bar{x}-x_0\|=\inf_{x\in\Omega}\|\bar{x}-x\|, \tag{3}$$

不等式

$$\langle\bar{x}-x_0,\ x-x_0\rangle\leqslant 0 \tag{4}$$

对于 $x\in\Omega$ 成立是表征 x_0 满足(3)的关系式.

证 设
$$d=\inf_{x\in\Omega}\|\bar{x}-x\|\geqslant 0,$$

$\{x_k\}$是(3)的极小化序列，即

$$x_k\in\Omega \qquad (k=1,\ 2,\ \cdots),$$

$$\lim_{k\to\infty}\|\bar{x}-x_k\|=d. \tag{5}$$

根据 $\mathbb{R}^n$ 中的平行四边形法则，成立着

$$\|\bar{x}-x_k\|^2+\|\bar{x}-x_l\|^2$$
$$=2\left\{\left\|\bar{x}-\frac{x_k+x_l}{2}\right\|^2+\left\|\frac{x_l-x_k}{2}\right\|^2\right\},$$

所以

$$\left\|\frac{x_l-x_k}{2}\right\|^2=\frac{1}{2}\{\|\bar{x}-x_k\|^2+\|\bar{x}-x_l\|^2\}-\left\|\bar{x}-\frac{x_k+x_l}{2}\right\|^2.$$

因为 Ω 是凸集，所以 $(x_k+x_l)/2\in\Omega$，从而上式末项不小于 d^2。

因此 $$0\leqslant\left\|\frac{x_k-x_l}{2}\right\|^2\leqslant\frac{1}{2}\{\|\bar{x}-x_k\|^2+\|\bar{x}-x_l\|^2\}-d^2.$$

上式右端当 $k, l\to\infty$ 时趋于 0，所以

$$\lim_{\substack{k\to\infty\\ l\to\infty}}\left\|\frac{x_k-x_l}{2}\right\|^2=0.$$

从而 $\{x_k\}$ 的极限 x_0 存在，由 Ω 的闭性推知 $x_0\in\Omega$。

因此 $$d=\lim_{k\to\infty}\|\bar{x}-x_k\|=\|\bar{x}-x_0\|.$$

即存在 x_0 使得 $\|\bar{x}-x_0\|=d$。

其次，证明唯一性．设 $\bar{x}_0\in\Omega$ 也满足

$$\|\bar{x}_0-x_0\|=d.$$

由平行四边形法则得到

$$\|\bar{x}-x_0\|^2+\|\bar{x}-\bar{x}_0\|^2$$
$$=2\left\{\left\|\bar{x}-\frac{x_0+\bar{x}_0}{2}\right\|^2+\left\|\frac{\bar{x}_0-x_0}{2}\right\|^2\right\}.$$

因此 $$d^2+d^2\geqslant 2d^2+\frac{1}{2}\|\bar{x}_0-x_0\|^2.$$

所以 $$\|\bar{x}_0-x_0\|^2\leqslant 0,$$

即 $\bar{x}_0=x_0$。

再证明不等式(4)．设 $x\in\Omega$，$\alpha\in(0, 1)$，那末

$$x_0+\alpha(x-x_0)=\alpha x+(1-\alpha)x_0\in\Omega,$$

从而

$$\|\bar{x}-x_0\|^2=d^2\leqslant\|\bar{x}-x_0-\alpha(x-x_0)\|^2$$

$$=\|\bar{x}-x_0\|^2-2\alpha\langle\bar{x}-x_0,\ x-x_0\rangle+\alpha^2\|x-x_0\|^2.$$

因此
$$2\alpha\langle\bar{x}-x_0,\ x-x_0\rangle\leqslant\alpha^2\|x-x_0\|^2.$$

两端除以 2α，并令 $\alpha\to0$，得到不等式(4)。

最后，设不等式(4)对于 $x\in\Omega$ 皆成立，那末

$$\|\bar{x}-x\|^2=\|\bar{x}-x_0-(x-x_0)\|^2$$

$$=\|\bar{x}-x_0\|^2-2\langle\bar{x}-x_0,\ x-x_0\rangle+\|x-x_0\|^2$$

$$\geqslant\|\bar{x}-x_0\|^2.$$

从而
$$\|\bar{x}-x_0\|^2\leqslant\inf_{x\in\Omega}\|\bar{x}-x\|^2,$$

即 x_0 适合(3)。

有时称不等式(4)为变分不等式。定理 3 表明：设 Ω 是凸闭集，那末对于 $\bar{x}\in\mathbb{R}^n$，存在唯一的 $x_0\in\Omega$ 适合变分不等式(4)。

定理 3 设 Ω 是 $\mathbb{R}^n$ 中的凸闭集，$\bar{x}\bar{\in}\Omega$，那末存在实数 c 和单位向量 $\lambda\in\mathbb{R}^n$，$\|\lambda\|=1$，使得超平面

$$\langle\lambda,\ x\rangle=c$$

把凸集 Ω 和 $\bar{x}$ 分在两侧，即

$$\langle\lambda,\ \bar{x}\rangle>c$$

而当 $x\in\Omega$ 时

$$\langle\lambda,\ x\rangle\leqslant c.$$

证 设 x_0 满足：$x_0\in\Omega$，

$$0<\|\bar{x}-x_0\|=\inf_{x\in\Omega}\|\bar{x}-x\|.$$

置
$$\lambda=(\bar{x}-x_0)/\|\bar{x}-x_0\|,\qquad c=\langle\lambda,\ x_0\rangle$$

那末
$$\|\lambda\|=1.$$

由不等式(4)得到：不等式

$$\langle\lambda,\ x-x_0\rangle\leqslant0$$

对 $x\in\Omega$ 成立，即当 $x\in\Omega$ 时

$$\langle \lambda,\ x\rangle \leqslant \langle \lambda,\ x_0\rangle = c.$$

又
$$\langle \lambda,\ \bar{x}-x_0\rangle = \|\bar{x}-x_0\| > 0,$$

所以
$$\langle \lambda,\ \bar{x}\rangle > \langle \lambda,\ x_0\rangle = c.$$

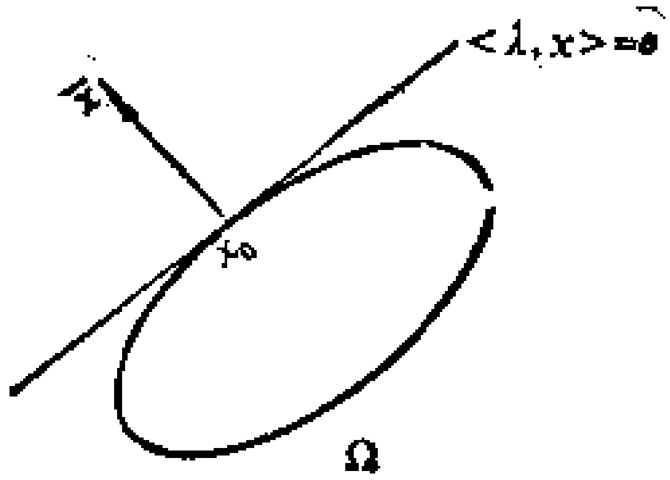


图 2.1

定理 4 设 Ω_1 和 Ω_2 是 $\mathbb{R}^n$ 中两个不相交的凸闭集，Ω_1 或 Ω_2 是有界的，那末存在 $c\in\mathbb{R}$，$\lambda\in\mathbb{R}^n$，$\|\lambda\|=1$，使得超平面

$$\langle \lambda,\ x\rangle = c$$

把 Ω_1 和 Ω_2 严格分在两侧，即成立着

$$\langle \lambda,\ x_1\rangle \geqslant c, \qquad \forall x_1\in\Omega_1,$$

$$\langle \lambda,\ x_2\rangle > c, \qquad \forall x_2\in\Omega_2.$$

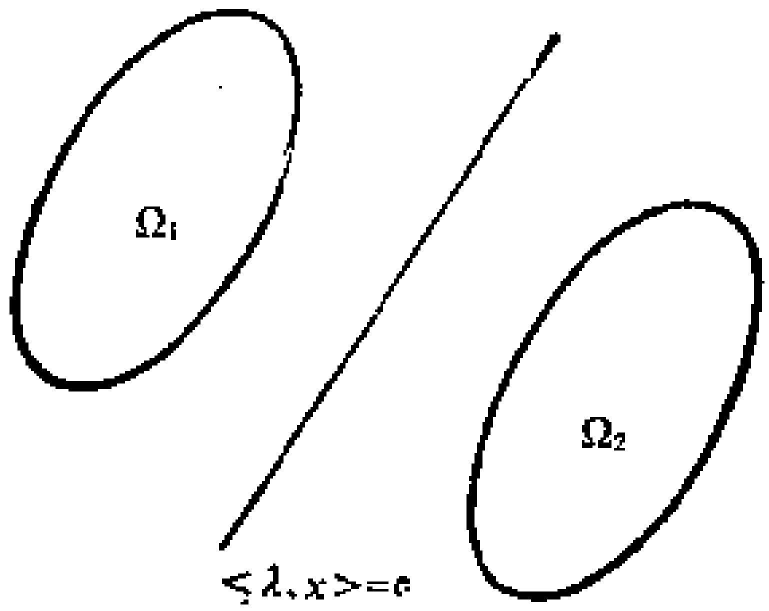


图 2.2

证 置
$$\Omega = \Omega_1 - \Omega_2,$$

那末根据定理 1 知 Ω 是凸闭集．因为 Ω_1 和 Ω_2 是不相交的，

所以 $0\bar{\in}\Omega$.

根据定理 3, 存在超平面

$$\langle\lambda, x\rangle=c_1$$

把 0 和 Ω 分为两侧, 即成立着

$$\langle\lambda, 0\rangle>c_1,$$

$$\langle\lambda, x\rangle\leqslant c_1, \quad \forall x\in\Omega.$$

因此, 成立着

$$c_1<0,$$

$$\langle\lambda, x_1-x_2\rangle\leqslant c_1, \quad \forall x_1\in\Omega_1, x_2\in\Omega_2.$$

从而当 $x_1\in\Omega_1$ 和 $x_2\in\Omega_2$ 时

$$\langle\lambda, x_1\rangle\leqslant c_1+\langle\lambda, x_2\rangle$$

置
$$c=\sup_{x_1\in\Omega_1}\langle\lambda, x_1\rangle,$$

那末
$$\langle\lambda, x_1\rangle\leqslant c\leqslant c_1+\langle\lambda, x_2\rangle.$$

从而
$$\langle\lambda, x_1\rangle\leqslant c, \quad \forall x_1\in\Omega_1,$$

$$\langle\lambda, x_2\rangle\geqslant c-c_1>c, \quad \forall x_2\in\Omega_2. \quad \blacksquare$$

定理 5 设 Ω 是 $\mathbb{R}^n$ 中的凸集, $\bar{x}$ 是 Ω 的边界点(即 $\bar{x}$ 的任一邻域中既有 Ω 的点又有不属于 Ω 的点), 那末存在过 $\bar{x}$ 的超平面

$$\langle\lambda, x\rangle=c,$$

使得 Ω 位在它的一侧, 即

$$\langle\lambda, \bar{x}\rangle=c,$$

$$\langle\lambda, x\rangle\leqslant c, \quad \forall x\in\Omega.$$

证 设 Ω 的极限点全体为 $\bar{\Omega}$, 即 $\bar{\Omega}$ 是 Ω 的闭包, 它是凸闭集. 显然, $\bar{x}$ 也是 $\bar{\Omega}$ 的边界点. 存在 $x_k\bar{\in}\bar{\Omega}$, 使得

$$\|\bar{\Omega}-x_k\|<1/k. \tag{6}$$

由于 $x_k\bar{\in}\bar{\Omega}$, 对 x_k 和 $\bar{\Omega}$ 应用定理 3 知道: 存在 $\lambda_k\in\mathbb{R}^n$, $\|\lambda_k\|=1$, 使得当 $x\in\bar{\Omega}$ 时

$$\langle\lambda_k, x\rangle<\langle\lambda_k, x_k\rangle. \tag{7}$$

根据 $\|\lambda_k\|=1$, 知 $\{\lambda_k\}$ 存在收敛子列,不妨设它就是 $\{\lambda_k\}$. 设它的极限为 $\lambda\in\mathbb{R}^n$, 且 $\|\lambda\|=1$.

在(7)中令 $k\to\infty$, 就得到

$$\langle\lambda, x\rangle\leqslant\langle\lambda, \bar{x}\rangle$$

对于 $x\in\bar{\Omega}$ 成立. ∎

定义 2 对于给定的 $c\in\mathbb{R}$ 和 $\lambda\in\mathbb{R}^n(\|\lambda\|=1)$, 满足不等式

$$\langle\lambda, x\rangle\leqslant c$$

的 x 的全体,称为闭的半空间.

显然,闭的半空间是凸闭集,但它是无界的.

定义 3 有限个闭的半空间的交集,如果含有内点,称为凸多面体.

凸多面体显然是凸闭集. 它可能是无界的,也可能是有界的. 集

$$\Omega=\{x\,|\,\langle\lambda_j, x\rangle\leqslant c_j, \qquad j=1, 2, \cdots, l\}$$

是有限个闭的半空间

$$\langle\lambda_j, x\rangle\leqslant c_j \qquad (j=1, 2, \cdots, l)$$

的交集.

如果存在超平面

$$\langle\lambda_0, x\rangle=c_0, \tag{8}$$

使得 Ω 位于其上,那末 Ω 不含有内点. 但可以在超平面(8)中讨论 Ω. Ω 可能有相对于(8)的内点. 因此,我们今后仅讨论含有内点的情形,即限于凸多面体的情形.

定义 4 设 Ω 是 $\mathbb{R}^n$ 中的集, 含有 Ω 的最小凸集称为 Ω 的凸包,记为 $\mathrm{co}\,\Omega$.

$\mathbb{R}^n$ 是含有 Ω 的凸集, 所有包含 Ω 的凸集的交集就是 Ω 的凸包. 所以上述定义是合理的.

定理 6 设 Ω 是 $\mathbb{R}^n$ 中的集, $x\in\mathrm{co}\,\Omega$, 那末存在 $n+1$ 个点

$x_j\in\Omega(j=0, 1, \cdots, n)$和$n+1$个实数$\alpha_j\geqslant 0$满足

$$\alpha_0x_0+\alpha_1x_1+\cdots+\alpha_nx_n=x, \tag{9}$$

$$\alpha_0+\alpha_1+\cdots+\alpha_n=1. \tag{10}$$

如果点x能用$x_0, x_1, \cdots, x_n$表示为(9)，而α_j适合关系

$$\alpha_j\geqslant 0, \quad \sum_{j=0}^{n}\alpha_j=1,$$

则称x是$x_0, x_1, \cdots, x_n$的凸组合．该定理表明：$\mathrm{co}\,\Omega$中的点是Ω中$n+1$个点的凸组合．

证 设

$$G=\left\{x\,\middle|\,x=\sum_{j=0}^{l}\alpha_jx_j,\ x_j\in\Omega,\ \alpha_j\geqslant 0,\ \sum_{j=0}^{l}\alpha_j=1,\ l=0, 1, \cdots\right\}.$$

显然，G包含Ω且是凸集．又，包含Ω的凸集必包含G．所以

$$G=\mathrm{co}\,\Omega.$$

如果$x\in\mathrm{co}\,\Omega$，那末存在$l+1$个点$x_j\in\Omega(j=0, 1, \cdots, l)$和$l+1$个实数$\alpha_j\geqslant 0(j=0, 1, \cdots, l)$满足

$$x=\alpha_0x_0+\alpha_1x_1+\cdots+\alpha_lx_l,$$

$$\alpha_0+\alpha_1+\cdots+\alpha_l=1. \tag{11}$$

定理要求证明$l\leqslant n$.

否则，设$l>n$，那末

$$x_0-x_l,\ x_1-x_l,\ \cdots,\ x_{l-1}-x_l$$

共有l个向量，它们在$\mathbb{R}^n$中是线性相关的，所以存在不全为0的实数$\beta_0, \beta_1, \cdots, \beta_{l-1}$，使得

$$\sum_{j=0}^{l-1}\beta_j(x_j-x_l)=0.$$

若记$\beta_l=-(\beta_0+\beta_1+\cdots+\beta_{l-1})$，那末

$$\beta_0+\beta_1+\cdots+\beta_l=0,$$

$$\beta_0x_0+\beta_1x_1+\cdots+\beta_lx_l=0. \tag{12}$$

我们不妨设(11)中的实数α_j都是正的(否则，可以丢掉该项).

现考虑 $$\frac{|\beta_0|}{\alpha_0}, \frac{|\beta_1|}{\alpha_1}, \cdots, \frac{|\beta_l|}{\alpha_l},$$

它们中必有最大者，不妨设它为 $|\beta_l|/\alpha_l$. 这时，成立着

$$|\beta_j|/\alpha_j \leqslant |\beta_l|/\alpha_l \quad (j=0, 1, \cdots, l-1).$$

所以

$$\left|\frac{\beta_j}{\beta_l}\right|\alpha_l \leqslant \alpha_j \quad (j=0, 1, \cdots, l-1). \tag{13}$$

用 α_l/β_l 乘(12)式，得到

$$\sum_{j=0}^{l-1}\beta_j\frac{\alpha_l}{\beta_l}x_j+\alpha_l x_l=0.$$

(9)减去上式，得到

$$\sum_{j=0}^{l-1}\left(\alpha_j-\beta_j\frac{\alpha_l}{\beta_l}\right)x_j=x. \tag{14}$$

由不等式(13)知道

$$\alpha_j-\beta_j\frac{\alpha_l}{\beta_l}\geqslant 0 \quad (j=0, 1, \cdots, l-1),$$

$$\sum_{j=0}^{l-1}\left(\alpha_j-\beta_j\frac{\alpha_l}{\beta_l}\right)=\sum_{j=0}^{l-1}\alpha_j-\sum_{j=0}^{l-1}\beta_j\alpha_l/\beta_l$$

$$=\sum_{j=0}^{l-1}\alpha_j+\alpha_l=1.$$

因此，由(14)知 x 可表示为 x_0、x_1、…、x_{l-1} 的凸组合. 如果 $l-1>n$，我们可以用同样手续. 因此，x 可表示为 Ω 中 $n+1$ 个点的凸组合. ∎

该定理是由卡拉提奥都里(C. Carathéodory)证明的.

定理 7 设 Ω 是 $\mathbb{R}^n$ 中的有界闭集，那末 $\mathrm{co}\Omega$ 是有界凸闭集.

证 因为 Ω 是有界的，所以存在常数 $M>0$，使得

$$\|x\| \leqslant M \quad \forall x \in \Omega.$$

对于 $x\in \mathrm{co}\,\Omega$，存在 $x_j\in\Omega$ 和 $\alpha_j\geqslant 0$ $(j=0, 1, \cdots, n)$，满足

$$x=\alpha_0x_0+\alpha_1x_1+\cdots+\alpha_nx_n,$$

$$\alpha_0+\alpha_1+\cdots+\alpha_n=1.$$

所以 $\|x\|\leqslant\alpha_0\|x_0\|+\alpha_1\|x_1\|+\cdots+\alpha_n\|x_n\|\leqslant M$.

即 $\mathrm{co}\,\Omega$ 是有界的.

再证 $\mathrm{co}\,\Omega$ 的闭性. 设 $\{x^{(k)}\}$ 收敛于 x_0, $x^{(k)}\in\mathrm{co}\,\Omega$. 根据定理 6, 对 $x^{(k)}\in\mathrm{co}\,\Omega$, 存在 $\alpha_j^{(k)}\geqslant 0$ 和 $x_j^{(k)}\in\Omega(j=0,1,\cdots,n)$, 满足

$$\alpha_0^{(k)}+\alpha_1^{(k)}+\cdots+\alpha_n^{(k)}=1,$$

$$\alpha_0^{(k)}x_0^{(k)}+\alpha_1^{(k)}x_1^{(k)}+\cdots+\alpha_n^{(k)}x_n^{(k)}=x^{(k)}.$$

因为 $\alpha_j^{(k)}$ 和 $x_j^{(k)}(k=1,2,\cdots)$ 都是有界的, 所以 $\{\alpha_j^{(k)}\}$ 和 $\{x_j^{(k)}\}$ $(j=0,1,\cdots,n)$ 都存在收敛子列, 不妨设是它们本身, 即存在下述极限

$$\lim_{k\to\infty}\alpha_j^{(k)}=\alpha_j,\qquad (j=0,1,\cdots,n),$$

$$\lim_{k\to\infty}x_j^{(k)}=x_j,\qquad (j=0,1,\cdots,n).$$

显然 $\alpha_j\geqslant 0(j=0,1,\cdots,n)$, $x_j\in\Omega(j=0,1,\cdots,n)$ (根据 Ω 的闭性). 并且

$$\alpha_0+\alpha_1+\cdots+\alpha_n=\lim_{k\to\infty}(\alpha_0^{(k)}+\alpha_1^{(k)}+\cdots+\alpha_n^{(k)})=1,$$

$$\sum_{j=0}^{n}\alpha_jx_j=\sum_{j=0}^{n}(\lim_{k\to\infty}\alpha_j^{(k)}\cdot\lim_{k\to\infty}x_j^{(k)})$$

$$=\lim_{k\to\infty}\sum_{j=0}^{n}\alpha_j^{(k)}x_j^{(k)}=\lim_{k\to\infty}x^{(k)}=x_0,$$

即 $x_0\in\mathrm{co}\,\Omega$, 所以 $\mathrm{co}\,\Omega$ 是闭的. ∎

定义 5 设 V 是 $\mathbb{R}^n$ 中的集, $x_0\in\mathbb{R}^n$. 如果对于 $x\in V$ 和 $\alpha>0$, 成立着 $x_0+\alpha(x-x_0)\in V$, 就称 V 是 $\mathbb{R}^n$ 中以 x_0 为顶点的锥. 如果 V 还是凸的, 就称 V 是凸锥.

例如, 半空间、超平面等都是凸锥. 又如以 x_0 为顶点的射线

$$L=\{x\mid x=x_0+\alpha(x_1-x_0),\ \alpha>0\}$$

也是凸锥, 这是 $x_1\in\mathbb{R}^n$, $x_1\neq x_0$.

显然, 当 V 是以 x_0 为顶点的锥时, $V-x_0$ 是以 0 为顶点的锥.

定理 8 设 V 是以 0 为顶点的凸锥, L 是以 0 为顶点的射线,

并且 L 的点都不是 V 的内点，那末存在 $\lambda\in\mathbb{R}^n$，$\|\lambda\|=1$，使得不等式

$$\langle\lambda,\ x\rangle\leqslant 0 \tag{15}$$

对于 $x\in V$ 成立；不等式

$$\langle\lambda,\ y\rangle\geqslant 0 \tag{16}$$

对于 $y\in L$ 成立.

证 设 $x_1\in L$，$x_1\neq 0$，那末

$$L=\{x\,|\,x=\alpha x_1,\ \alpha>0\}.$$

由于 x_1 不是凸锥 V 的内点，根据定理 3 和定理 5，存在 $c\in\mathbb{R}$ 和 $\lambda\in\mathbb{R}^n$，$\|\lambda\|=1$，使得超平面

$$\langle\lambda,\ x\rangle=c$$

把 V 和 x_1 分在两侧，即

$$\langle\lambda,\ x_1\rangle\geqslant c, \tag{17}$$

而当 $x\in V$ 时

$$\langle\lambda,\ x\rangle\leqslant c.$$

由于当 $x\in V$，$\mu>0$ 时 $\mu x\in V$，所以由上式得到

$$\mu\langle\lambda,\ x\rangle\leqslant c. \tag{18}$$

上式两端除以 μ，并令 $\mu\to+\infty$，得到

$$\langle\lambda,\ x\rangle\leqslant 0. \tag{19}$$

由(18)令 $\mu\to 0$，就得到

$$0\leqslant c.$$

从而由(17)得到

$$\langle\lambda,\ x_1\rangle\geqslant 0,$$

从而当 $y=\alpha x_1$，$\alpha>0$ 时，即当 $y\in L$ 时

$$\langle\lambda,\ y\rangle\geqslant 0. \tag{20}$$

因此，(15)和(16)成立. ∎

定理 9 设 Ω 是 $\mathbb{R}^n$ 中的有界的凸闭集，那末它是有限个或

可数个闭的半空间的交集.

证 首先，对于 $x_0 \bar{\in} \Omega$，根据定理 4，存在 $\lambda \in \mathbb{R}^n$，$\|\lambda\|=1$，使得

$$\max_{x\in\Omega}\langle\lambda, x\rangle<\langle\lambda, x_0\rangle.$$

其次，存在 $\{\lambda_k\}$，$\lambda_k\in\mathbb{R}^n$，$\|\lambda_k\|=1$，使得 $\{\lambda_k\}$ 在单位球面上是稠密的. 从而存在 λ_{k_0}，使得 $\|\lambda_{k_0}-\lambda\|<\delta$，这里

$$\delta=\{\langle\lambda, x_0\rangle-\max_{x\in\Omega}\langle\lambda, x\rangle\}/2\{\max_{x\in\Omega}\|x\|+\|x_0\|\}.$$

记 $M=\|x_0\|+\max\{\|x\| \,|\, x\in\Omega\}$，那末当 $x\in\Omega$ 时

$$\langle\lambda, x\rangle-\langle\lambda, x_0\rangle\leqslant-2M\delta.$$

从而

$$\begin{aligned}&\langle\lambda_{k_0}, x\rangle-\langle\lambda_{k_0}, x_0\rangle\\&\leqslant\langle\lambda, x\rangle-\langle\lambda, x_0\rangle+\delta(\|x\|+\|x_0\|)\\&\leqslant M\delta-2M\delta=-M\delta.\end{aligned}$$

即 Ω 位在半空间 $H_{k_0}=\{x \,|\, \langle\lambda_{k_0}, x-x_0\rangle\leqslant-M\delta\}$ 中. 从而

$$\Omega\subset\bigcap H_{k_0}.$$

显然，$x_0\bar{\in} H_{k_0}$，即 $x_0\bar{\in}\bigcap H_{k_0}$. 所以 $\Omega=\bigcap H_{k_0}$. ∎

习　　题

1　设 A 是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性有界算子，$b\in\mathbb{R}^n$，如果存在 x 适合

$$Ax=b,$$

试证：$H=\{x \,|\, Ax=b,\ x\in\mathbb{R}^n\}$ 是凸闭集.

2　设 H 如题 1，$\bar{x}\bar{\in} H$，试给出表征适合

$$\|\bar{x}-x_0\|=\min_{x\in H}\|\bar{x}-x\|$$

的 x_0 的关系式.

3　设 A 是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的自共轭线性有界算子，满足关系

$$\langle Ax, y\rangle=\langle x, Ay\rangle,\quad \forall x, y\in\mathbb{R}^n,$$
$$\langle Ax, x\rangle\geqslant\alpha\|x\|^2,\quad \forall x\in\mathbb{R}^n,$$

这里 $\alpha>0$ 是常数. 又设 $b\in\mathbb{R}^n$，Ω 是 $\mathbb{R}^n$ 中的凸闭集. 试讨论下述问题：

$$\inf_{x\in\Omega}\{\langle Ax, x\rangle-2\langle b, x\rangle\}.$$

4　试证凸锥的内部是凸锥.

§3 线性系统的等时区域

考虑线性系统

$$
\begin{aligned}
&\frac{dx(t)}{dt}=A(t)x(t)+B(t)u(t), \\
&x(t_0)=x_0,
\end{aligned}
\tag{1}
$$

这里 $x\in\mathbb{R}^n$ 是系统的状态向量，$u\in\mathbb{R}^m$ 是系统的控制向量，$A(\cdot)$: $[t_0, t_1]\to\mathbb{R}^{n\times n}$ 和 $B(\cdot)$: $[t_0, t_1]\to\mathbb{R}^{n\times m}$ 是可积的，$x_0\in\mathbb{R}^n$ 给定.

设 U 是 $\mathbb{R}^m$ 中的有界凸闭集，称为系统(1)的控制区域．它表征控制作用所受的约束条件．如果 $u(\cdot)$: $[t_0, t_1]\to\mathbb{R}^m$ 是可积的，并且 $u(t)\in U$ 几乎处处成立，就称 $u(\cdot)$是容许控制，记为

$$u(\cdot)\in U_{ad}.$$

当 $u(\cdot)$: $[t_0, t_1]\to U$ 是逐段连续函数时，它是容许控制，这时控制作用可以瞬时地从一状态变为另一状态．当 $u(\cdot)\in U_{ad}$ 给定时，(1)相应的解 $x=x(\cdot, u)$ 可表示为

$$x(t, u)=\Phi(t, t_0)x_0+\int_{t_0}^{t}\Phi(t, s)B(s)u(s)ds, \tag{2}$$

其中 Φ 是方程组

$$\frac{dx(t)}{dt}=A(t)x(t)$$

的转移矩阵，即 Φ 适合

$$\Phi(t, s)=I+\int_{s}^{t}A(\sigma)\Phi(\sigma, s)d\sigma.$$

定义 1 对于给定的 $t\in[t_0, t_1]$，集

$$R(t)\triangleq\{x\,|\,x=x(t, u), u(\cdot)\in U_{ad}\}$$

称为系统(1)的等时区域.

显然，$R(t)$ 是系统 (1) 在容许控制作用下于时刻 t 能达到的点的全体.

定理1 对于 $t\in[t_0, t_1]$，等时区域 $R(t)$ 是有界凸集.

证 由表达式(2)和 U 的有界性，容易知道 $R(t)$ 是有界的.

如果 $u_j(\cdot)\in U_{ad}(j=1, 2)$，那末当 $s\in[t_0, t_1]$ 时

$$u_j(s)\in U \qquad (j=1, 2).$$

根据 U 的凸性，推知：当 $\alpha\in[0, 1]$ 时

$$\alpha u_1(s)+(1-\alpha)u_2(s)\in U.$$

所以 $\alpha u_1(\cdot)+(1-\alpha)u_2(\cdot)\in U_{ad}$.

由表达式(2)得到

$$x(t, \alpha u_1+(1-\alpha)u_2)=\alpha x(t, u_1)+(1-\alpha)x(t, u_2),$$

即 $\alpha x(t, u_1)+(1-\alpha)x(t, u_2)\in R(t)$. 所以，$R(t)$是凸集. ▌

为讨论 $R(t)$的闭性，我们需要下面的引理.

引理1 设 U 是 $\mathbb{R}^m$ 中的有界集，$u_k(\cdot)\in U_{ad}$，即

$$u_k(t)\in U \qquad (k=1, 2, \cdots)$$

在 $[t_0, t_1]$ 上几乎处处成立. 那末 $\{u_k(\cdot)\}$ 存在弱 * 收敛的子列 $\{u_{k_j}(\cdot)\}$，即存在 $u(\cdot)$使得等式

$$\lim_{k_j\to\infty}\int_{t_0}^{t_1}\langle v(t), u_{k_j}(t)\rangle dt=\int_{t_0}^{t_1}\langle v(t), u(t)\rangle dt \tag{3}$$

对任一勒贝格可积的 $v(\cdot):[t_0, t_1]\to\mathbb{R}^m$ 成立.

证 设 $L=L([t_0, t_1]; \mathbb{R}^m)$ 表示可积的 $v(\cdot):[t_0, t_1]\to\mathbb{R}^m$ 的全体，且对于 $v(\cdot)\in L$ 赋以范数

$$\|v(\cdot)\|=\int_{t_0}^{t_1}\|v(t)\|dt.$$

根据第一章 § 1 习题 9 和 10 的结论，在 L 中的函数皆可以用分点为有理数，值为有理点的简单函数来任意逼近. 这些简单函数的全体是可列的，记之为 $\{v_j(\cdot)\}(j=1, 2, \cdots)$. 上述结论是说，对于 $v(\cdot)\in L$，以及正数 $\varepsilon>0$，存在某一个 $v_{j_0}(\cdot)$，满足

$$\|v(\cdot)-v_{j_0}(\cdot)\|<\varepsilon. \tag{4}$$

因为 $u_k(\cdot)\in U_{ad}(k=1, 2, \cdots)$, 所以

$$u_k(t)\in U \qquad (k=1, 2, \cdots).$$

但 U 是有界的，所以存在常数 $M>0$, 使得

$$\|u_k(t)\|\leqslant M \qquad (k=1, 2, \cdots) \tag{5}$$

对于 $t\in[t_0, t_1]$ 皆成立。

考察数列

$$\int_{t_0}^{t_1}\langle v_1(t), u_k(t)\rangle dt. \tag{6}$$

因为 $$\left|\int_{t_0}^{t_1}\langle v_1(t), u_k(t)\rangle dt\right|\leqslant M\int_{t_0}^{t_1}\|v_1(t)\|dt,$$

所以(6)是有界数列，从而存在收敛子列，即极限

$$\lim_{k\to\infty}\int_{t_0}^{t_1}\langle v_1(t), u_k^{(1)}(t)\rangle dt$$

存在，这里 $\{u_k^{(1)}(\cdot)\}\subset\{u_k(\cdot)\}$.

再考察数列

$$\int_{t_0}^{t_1}\langle v_2(t), u_k^{(1)}(t)\rangle dt,$$

它也是有界数列，存在收敛子列，设 $\{u_k^{(2)}\}\subset\{u_k^{(1)}\}$ 使得极限

$$\lim_{k\to\infty}\int_{t_0}^{t_1}\langle v_2(t), u_k^{(2)}(t)\rangle dt$$

存在、显然，这时极限

$$\lim_{k\to\infty}\int_{t_0}^{t_1}\langle v_1(t), u_k^{(2)}(t)\rangle dt$$

也存在，这是因为 $\{u_k^{(2)}(\cdot)\}\subset\{u_k^{(1)}(\cdot)\}$. 逐次进行下去，得到 $\{u_k(\cdot)\}$ 的子列 $\{u_k^{(l)}(\cdot)\}\subset\{u_k^{(l-1)}(\cdot)\}\subset\cdots\subset\{u_k(\cdot)\}$，使得极限

$$\lim_{k\to\infty}\int_{t_0}^{t_1}\langle v_j(t), u_k^{(l)}(t)\rangle dt \qquad (j=1, 2, \cdots, l)$$

存在. 取对角线序列 $\{u_l^{(l)}(\cdot)\}$，它是 $\{u_k(\cdot)\}$的子列，并且对于任何给定的 j, 极限

$$\lim_{l\to\infty}\int_{t_0}^{t_1}\langle v_j(t),\ u_l^{(l)}(t)\rangle dt$$

存在，这是因为：除有限项外，序列 $\{u_l^{(l)}(\cdot)\}$ 是 $\{u_k^{(j)}(\cdot)\}$ 的子列．

现在证明：对于 $v(\cdot)\in L([t_0,\ t_1];\ \mathbb{R}^m)$，极限

$$\lim_{l\to\infty}\int_{t_0}^{t_1}\langle v(t),\ u_l^{(l)}(t)\rangle dt \tag{7}$$

是存在的．事实上，因为

$$\begin{aligned}&\left|\int_{t_0}^{t_1}\langle v(t),\ u_l^{(l)}(t)-u_q^{(q)}(t)\rangle dt\right|\\ &\leqslant\left|\int_{t_0}^{t_1}\langle v(t)-v_j(t),\ u_l^{(l)}(t)-u_q^{(q)}(t)\rangle dt\right|\\ &\quad+\left|\int_{t_0}^{t_1}\langle v_j(t),\ u_l^{(l)}(t)-u_q^{(q)}(t)\rangle dt\right|.\end{aligned}$$

对于 $\varepsilon>0$，取 $v_j(\cdot)$ 满足

$$\|v(\cdot)-v_j(\cdot)\|<\frac{\varepsilon}{3M},$$

对于 $v_j(\cdot)$，存在 N，使得当 $l,\ q\geqslant N$ 时

$$\left|\int_{t_0}^{t_1}\langle v_j(t),\ u_l^{(l)}(t)-u_q^{(q)}(t)\rangle dt\right|<\frac{\varepsilon}{3}.$$

这时，当 $l,\ q\geqslant N$ 时

$$\left|\int_{t_0}^{t_1}\langle v(t),\ u_l^{(l)}(t)-u_q^{(q)}(t)\rangle dt\right|<\varepsilon.$$

所以极限(7)是存在的．记(7)为 $\alpha(v(\cdot))$．显然，成立着

$$|\alpha(v(\cdot))|\leqslant M\int_{t_0}^{t_1}\|v(t)\|dt.$$

并且 $\alpha(v(\cdot))$ 在 L 上是线性的．即得到：$\alpha(v(\cdot))$ 是空间 $L([t_0,t_1];\ \mathbb{R}^m)$ 上的线性连续泛函．

根据 $L([t_0,\ t_1];\ \mathbb{R}^m)$ 上线性连续泛函的表示定理(见附录 1)，存在有界可积的 $u(\cdot):[t_0,\ t_1]\to\mathbb{R}^m$，使得

$$\alpha(v(\cdot))=\int_{t_0}^{t_1}\langle v(t),\ u(t)\rangle dt.$$

这就是说，成立着

$$\lim_{l\to+\infty}\int_{t_0}^{t_1}\langle v(t),\ u_l^{(l)}(t)\rangle dt=\int_{t_0}^{t_1}\langle v(t),\ u(t)\rangle dt.$$

所以，序列 $\{u_l^{(l)}(\cdot)\}$ 弱*收敛于 $u(\cdot)$. ▮

引理 2 设 U 是 $\mathbb{R}^m$ 中的有界凸闭集，$u_k(\cdot)\in U_{ad}(k=1,2,\cdots)$，并且 $\{u_k(\cdot)\}$ 弱*收敛于 $u(\cdot)$，即对于 $v(\cdot)\in L([t_0,t_1];\mathbb{R}^m)$，成立着

$$\lim_{k\to+\infty}\int_{t_0}^{t_1}\langle v(t),\ u_k(t)\rangle dt=\int_{t_0}^{t_1}\langle v(t),\ u(t)\rangle dt, \tag{8}$$

那末 $u(\cdot)\in U_{ad}$.

证 根据 § 2 定理 9, U 是有限个或可数个闭的半空间的交集．设

$$\langle\lambda,\ u\rangle\leqslant c$$

是上述闭的半空间中的一个．

由假设 $u_k(\cdot)\in U_{ad}(k=1,2,\cdots)$，所以，除去一个零集外，在 $[t_0,t_1]$ 上成立着

$$\langle\lambda,\ u_k(t)\rangle\leqslant c\qquad(k=1,\ 2,\ \cdots).$$

现在证明：不等式

$$\langle\lambda,\ u(t)\rangle\leqslant c$$

在 $[t_0,t_1]$ 上除一零集外成立．

否则，集 $E=\{t\in[t_0,t_1]\,|\,\langle\lambda,\ u(t)\rangle>c\}$

的测度 $\mu(E)>0$. 置

$$v(t)=\begin{cases}\lambda, & 当\ t\in E\ 时,\\ 0, & 当\ t\in[t_0,t_1]\setminus E\ 时,\end{cases}$$

即
$$v(t)=\lambda\chi_E(t).$$

那末 $v(\cdot):[t_0,t_1]\to\mathbb{R}^m$ 是可积的．根据(8)，成立着

$$\lim_{k\to\infty}\int_{t_0}^{t_1}\langle\lambda,\ u_k(t)\rangle\chi_E(t)dt$$

$$=\int_{t_0}^{t_1}\langle\lambda,\ u(t)\rangle\chi_E(t)dt=\int_E\langle\lambda,\ u(t)\rangle dt. \tag{9}$$

上式右端大于 $\mu(E)\cdot c$.

但是 $$\int_{t_0}^{t_1}\langle\lambda,\ u_k(t)\rangle\chi_E(t)dt=\int_E\langle\lambda,\ u_k(t)\rangle dt$$

$$\leqslant c\cdot\mu(E),$$

所以 $$\lim_{k\to+\infty}\int_{t_0}^{t_1}\langle\lambda,\ u_k(t)\rangle\chi_E(t)dt\leqslant c\cdot\mu(E).$$

因此，由(5)式及上式得到

$$\int_E\langle\lambda,\ u(t)\rangle dt\leqslant c\cdot\mu(E).$$

但已证 $$\int_E\langle\lambda,\ u(t)\rangle dt>c\mu(E).$$

从而得到矛盾不等式

$$c\mu(E)<c\mu(E).$$

因此，不等式

$$\langle\lambda,\ u(t)\rangle\leqslant c \tag{10}$$

在 $[t_0,\ t_1]$ 上几乎处处成立.

因为有界凸闭集 U 是有限个或可数个闭的半空间的交集，所以

$$u(t)\in U$$

在 $[t_0,\ t_1]$ 上几乎处处成立. ▌

定理 2 对于给定的 $t\in[t_0,\ t_1]$, $R(t)$ 是闭集.

证 设 $x_k\in R(t)$, $\lim\limits_{k\to\infty}x_k=\bar{x}$. 我们要证 $\bar{x}\in R(t)$.

因为 $x_k\in R(t)$, 所以存在 $u_k(\cdot)\in U_{ad}$, 使得

$$x_k=\Phi(t,\ t_0)x_0+\int_{t_0}^{t}\Phi(t,\ s)B(s)u_k(s)ds. \tag{11}$$

根据引理 1, $\{u_k(\cdot)\}$ 存在弱* 收敛的子列，仍记该子列为

$\{u_k(\cdot)\}$. 根据引理 2, $\{u_k(\cdot)\}$的弱* 极限 $u(\cdot)\in U_{ad}$. 在 (11) 中令 $k\to\infty$, 得到

$$\bar{x}=\Phi(t,\ t_0)x_0+\int_{t_0}^{t}\Phi(t,\ s)B(s)u(s)ds,$$

即 $\bar{x}\in R(t)$. ∎

对于 $\mathbb{R}^n$ 中的两个集 P 和 Q, 我们规定它们之间的豪斯道夫 (F. Hausdorff) 距离 $d(P,\ Q)$ 为

$$d(P,\ Q)=\frac{1}{2}\{\sup_{y\in P}\inf_{x\in Q}\|x-y\|+\sup_{x\in Q}\inf_{y\in P}\|x-y\|\}. \tag{12}$$

定理 3 集 $R(t)$ 关于 $t\in[t_0,\ t_1]$ 是连续变动的, 即

$$\lim_{t\to\bar{t}}d(R(t),\ R(\bar{t}))=0.$$

证 由表达式 (2) 得到

$$\begin{aligned}x(t,\ u)-x(\bar{t},\ u)=&\{\Phi(t,\ t_0)x_0-\Phi(\bar{t},\ t_0)x_0\}\\&+\int_{t_0}^{\bar{t}}\{\Phi(t,\ s)-\Phi(\bar{t},\ s)\}B(s)u(s)ds\\&+\int_{\bar{t}}^{t}\Phi(t,\ s)B(s)u(s)ds.\end{aligned}$$

因为存在常数 $M>0$, 使得

$$\|u(s)\|\leqslant M,$$

所以

$$\begin{aligned}\|x(t,\ u)-x(\bar{t},\ u)\|\leqslant&\|\Phi(t,\ t_0)-\Phi(\bar{t},\ t_0)\|\|x_0\|\\&+\int_{t_0}^{\bar{t}}\|\Phi(t,\ s)-\Phi(\bar{t},\ s)\|\|B(s)\|Mds\\&+\int_{\bar{t}}^{t}\|\Phi(t,\ s)\|\|B(s)\|Mds.\end{aligned}$$

因此, $x(t,\ u)$ 与 $R(\bar{t})$ 的距离不大于上式的右端; 同样, $x(\bar{t},\ u)$ 与 $R(t)$ 的距离不大于上式的右端. 从而, $R(t)$ 与 $R(\bar{t})$ 之间的豪斯道夫距离不大于上式的右端. 但当 $t\to\bar{t}$ 时, 上式右端趋于 0. 从而得证定理. ∎

在定理 3 的证明中, 仅要求 U 是有界集, 并不要求 U 是凸闭

集. 从定理1和定理2的证明看到, U 的凸性是十分重要的条件.

然而, 从绪论的例题看, 如果要求容许控制是取值1, -1 和0的可积函数, 那末

$$U=\{1,\ 0,\ -1\} \tag{13}$$

并不具有凸性. 当然, 如果扩充要求容许控制可取区间 $[-1,\ 1]$ 中的任一点, 那末

$$\tilde{U}=[-1,\ 1]$$

是凸集, 且 $\tilde{U}=\mathrm{co}\,U$.

我们讨论等时区域的凸闭性, 是为了以后研究控制问题. 从控制问题本身来说, 一般不应当要求 U 是凸的. 特别, 布绍在讨论最优开关控制时, 假设 U 仅为上面所示的集(13). 因此, 为了数学结果能用于控制问题的实际, 有必要探讨 U 不具凸性条件时的结论. 下面将指出, 当 U 是有界闭集时, $R(t)$ 仍是凸闭集.

我们首先证明

定理4 设 U 是 $\mathbb{R}^m$ 的任一有界集, 那末 $R(t)$ 的闭包 $\overline{R}(t)$ 是凸的.

证 设 x_1、$x_2\in\overline{R}(t)$, $\alpha\in(0,\ 1)$, 要证明

$$\alpha x_1+(1-\alpha)x_2\in\overline{R}(t).$$

对于 $\varepsilon>0$, 存在 $u_j(\cdot)\in U_{ad}(j=1,\ 2)$ 使得

$$\|x_j-x(t,\ u_j)\|<\varepsilon/2 \qquad (j=1,\ 2). \tag{14}$$

在区间 $[t_0,\ t]$ 上考虑

$$\Phi(t,\ \cdot)B(\cdot)\{u_1(\cdot)-u_2(\cdot)\},$$

它是可积的, 根据定理 I-1-3①(第一章§1定理3), 存在可测集 $E_\alpha(\varepsilon)\subset[t_0,\ t]$, 满足

$$\mu(E_\alpha(\varepsilon))=\alpha(t-t_0),$$

① 定理 I-1-3 表示第一章§1定理3.

$$\alpha\int_{t_0}^{t}\Phi(t,s)B(s)\{u_1(s)-u_2(s)\}ds$$

$$=\int_{E_\alpha(\varepsilon)}\Phi(t,s)B(s)\{u_1(s)-u_2(s)\}ds+\eta,$$

$$\|\eta\|<\varepsilon/2. \tag{15}$$

因此

$$\alpha x(t,u_1)+(1-\alpha)x(t,u_2)$$

$$=\Phi(t,t_0)x_0+\int_{t_0}^{t}\Phi(t,s)B(s)u_2(s)ds$$

$$+\alpha\int_{t_0}^{t}\Phi(t,s)B(s)\{u_1(s)-u_2(s)\}ds$$

$$=\Phi(t,t_0)x_0+\int_{t_0}^{t}\Phi(t,s)B(s)u_2(s)ds$$

$$+\int_{E_\alpha(\varepsilon)}\Phi(t,s)B(s)\{u_1(s)-u_2(s)\}ds+\eta.$$

置 $$u(s)=\begin{cases}u_1(s), & \text{当 } s\in E_\alpha(\varepsilon) \text{ 时},\\ u_2(s), & \text{当 } s\in[t_0,t]\setminus E_\alpha(\varepsilon) \text{ 时},\end{cases}$$

那末 $u(s)\in U$，从而 $u(\cdot)\in U_{ad}$，并且

$$\alpha x(t,u_1)+(1-\alpha)x(t,u_2)$$

$$=\Phi(t,t_0)x_0+\int_{t_0}^{t}\Phi(t,s)B(s)u(s)ds+\eta,$$

即 $$\alpha x(t,u_1)+(1-\alpha)x(t,u_2)=x(t,u)+\eta.$$

根据上式和(14)，(15)，得到

$$\|\alpha x_1+(1-\alpha)x_2-x(t,u)\|<\varepsilon.$$

即 $\alpha x_1+(1-\alpha)x_2\in\overline{R}(t)$。 ∎

定理 5 设 $y(\cdot):[\alpha,\beta]\to R^n$ 和 $\lambda(\cdot):[\alpha,\beta]\to[0,1]$ 是可积的，那末存在可测集 $E\subset[\alpha,\beta]$，使得

$$\int_\alpha^\beta\lambda(t)y(t)dt=\int_E y(t)dt.$$

该定理是李雅普诺夫 (A. A. Ляпунов) 于 1940 年证明的，称

为向量值测度的值域定理. 原来的证明很长，后来有许多简化证明. 下面的证明是基于凸集的端点理论的. 下述引理是克莱因(М.Г.Крейн)-米尔曼(Д. П. Мильман)定理的特例.

引理 3 设 Ω 是线性赋范空间的共轭空间中的有界凸闭集，那末 Ω 存在端点，即存在一点属于 Ω，但它又不是 Ω 中线段的中点. 并且端点集的闭凸包就是 Ω.

该引理所述的概念、结论和证明，读者可以在关肇直、张恭庆、冯德兴著«线性泛函分析入门»[5]一书中找到，这里省略了.

下面是雍炯敏关于定理 5 的证明(«数学研究与评论»〈4〉①，1984 年第 2 期第 4 页).

设
$$a=\int_\alpha^\beta \lambda(t)y(t)dt.$$

又设 Ω 是使
$$\int_\alpha^\beta \tilde{\lambda}(t)y(t)dt=a$$
成立的可积的 $\tilde{\lambda}(\cdot):[\alpha,\beta]\to[0,1]$ 的全体. 显然，$\tilde{\lambda}(\cdot)$ 是 $[\alpha,\beta]$ 上有界可积函数空间 $L^\infty([\alpha,\beta])$ 的点. Ω 是 L^∞ 中的凸集. 利用引理 1，还可以证明 Ω 是 L^∞ 中的闭集.

根据 L^∞ 是 L 的共轭空间，而 L 是线性赋范空间，所以由引理 3 知：Ω 存在端点 $\lambda_0(\cdot):[\alpha,\beta]\to[0,1]$，使得
$$\int_\alpha^\beta \lambda_0(t)y(t)dt=a.$$

现在证明 $\lambda_0(\cdot)$ 是可测集的特征函数，即证明
$$\mu\{t|0<\lambda_0(t)<1\}=0.$$

否则，存在 $\varepsilon>0$ 和正测度集 E，使得不等式
$$\varepsilon\leqslant\lambda_0(t)\leqslant 1-\varepsilon$$
对于 $t\in E$ 成立.

---

① 〈4〉表示参考文献序号。

因为 $\mu(E)>0$，所以存在 $E_1, E_2, \cdots, E_{n+1}$ 满足

$$E=\bigcup_{j=1}^{n+1} E_j,$$

$$E_j\cap E_k=\varnothing \qquad (j\neq k),$$

$$\mu(E_j)>0.$$

考察 $$x_j=\int_{E_j} y(t)dt \qquad (j=1, 2, \cdots, n+1),$$

它们在 $\mathbb{R}^n$ 中是线性相关的，所以存在不全为 0 的数 $c_1, c_2, \cdots, c_{n+1}$，使得

$$c_1x_1+c_2x_2+\cdots+c_{n+1}x_{n+1}=0,$$

并且 $$|c_j|<\varepsilon \qquad (j=1, 2, \cdots, n+1).$$

置 $$h(t)=c_1\chi_{E_1}(t)+c_2\chi_{E_2}(t)+\cdots+c_{n+1}\chi_{E_{n+1}}(t),$$

那末 $$|h(t)|<\varepsilon,$$

并且 $$\int_\alpha^\beta h(t)y(t)dt=\sum_{j=1}^{n+1} c_j\int_\alpha^\beta \chi_{E_j}(t)y(t)dt$$

$$=\sum_{j=1}^{n+1} c_j\int_{E_j} y(t)dt=0.$$

这样，成立着

$$\int_\alpha^\beta \{\lambda_0(t)\pm h(t)\}y(t)dt=a,$$

并且当 $t\in[\alpha, \beta]$ 时

$$0\leqslant\lambda_0(t)\pm h(t)\leqslant 1,$$

即 $\lambda_0(\cdot)+h(\cdot)$ 和 $\lambda_0(\cdot)-h(\cdot)$ 都是 Ω 的点，而 $\lambda_0(\cdot)$ 是这两点联线的中点. 它与 $\lambda_0(\cdot)$ 为 Ω 的端点假设相矛盾.

因此，$\lambda_0(\cdot)$ 是某一可测集 E 的特征函数，即

$$\int_\alpha^\beta \lambda(t)y(t)dt=\int_\alpha^\beta \lambda_0(t)y(t)dt=\int_E y(t)dt. \qquad ∎$$

定理 6 设 U 是 $\mathbb{R}^m$ 中的任一集（可能是无界的），U_{ad} 是取值 U 的有界可积函数的全体，那末

$$R(t)=\{x|x=x(t,\ u),\ u(\cdot)\in U_{ad}\}$$

是 $\mathbb{R}^n$ 中的凸集.

证 设 $\alpha\in(0,\ 1)$，并且

$$x_j=\Phi(t,\ t_0)x_0+\int_{t_0}^{t}\Phi(t,\ s)B(s)u_j(s)ds,$$

$$u_j(\cdot)\in U_{ad}\quad(j=1,\ 2),$$

那末
$$\alpha x_1+(1-\alpha)x_2=\Phi(t,\ t_0)x_0$$
$$+\int_{t_0}^{t}\Phi(t,\ s)B(s)u_2(s)ds$$
$$+\alpha\int_{t_0}^{t}\Phi(t,\ s)B(s)\{u_1(s)-u_2(s)\}ds.$$

根据李雅普诺夫定理，存在可测集 $E\subset[t_0,\ t]$，使得上式右端的末项等于

$$\int_E\Phi(t,\ s)B(s)\{u_1(s)-u_2(s)\}ds.$$

置
$$u(s)=\begin{cases}u_1(s), & 当\ s\in E\ 时,\\ u_2(s), & 当\ s\in[t_0,\ t]\backslash E\ 时,\end{cases}$$

那末 $u(\cdot)\in U_{ad}$，并且

$$\alpha x_1+(1-\alpha)x_2=\Phi(t,\ t_0)x_0$$
$$+\int_{t_0}^{t}\Phi(t,\ s)B(s)u(s)ds,$$

即
$$\alpha x_1+(1-\alpha)x_2\in R(t),$$

所以 $R(t)$ 是凸集. ▮

注 1 李雅普诺夫定理可以改写为：设 E 是 $[\alpha,\ \beta]$ 中的可测集，$y(\cdot):E\to\mathbb{R}^n$ 和 $\lambda(\cdot):E\to[0,\ 1]$ 是可积的，那末存在可测集 $E_1\subset E$，使得成立着

$$\int_E\lambda(t)y(t)dt=\int_{E_1}y(t)dt. \tag{16}$$

事实上，在 $[\alpha,\ \beta]$ 上考虑 $y(t)\chi_E(t)$ 和 $\lambda(t)\chi_E(t)$，那末存在

$\widetilde{E}\subset[\alpha, \beta]$，使得

$$\int_\alpha^\beta \lambda(t)\chi_E(t)y(t)dt=\int_{\widetilde{E}}\chi_E(t)y(t)dt,$$

置 $E_1=E\cap\widetilde{E}$，就得到 (16)。

定理 7 设 $E\subset[\alpha, \beta]$ 是可测集，$y_j(\cdot):E\to\mathbb{R}^n$ 和 $\lambda_j(\cdot):E\to[0, 1]$ 是可积的，$j=1, 2, \cdots, l$，$\lambda_1(t)+\lambda_2(t)+\cdots+\lambda_l(t)=1$ 在 E 上几乎处处成立，那末存在 $E_j\subset E(j=1, 2, \cdots, l)$，使得

$$E=E_1\cup E_2\cup\cdots\cup E_l,$$

$$E_j\cap E_k=\varnothing \qquad (j\neq k),$$

$$\sum_{j=1}^l\int_E\lambda_j(t)y_j(t)dt=\sum_{j=1}^l\int_{E_j}y_j(t)dt. \tag{17}$$

证 当 $l=1$ 时，上式自然成立。

当 $l=2$ 时，上式左端等于

$$\int_E y_2(t)dt+\int_E\lambda_1(t)\{y_1(t)-y_2(t)\}dt,$$

根据引理 3 的注 1，存在 $E_1\subset E$，使得上式末项等于

$$\int_{E_1}\{y_1(t)-y_2(t)\}dt.$$

因此 $$\int_E\lambda_1(t)y_1(t)dt+\int_E\lambda_2(t)y_2(t)dt$$

$$=\int_{E_1}y_1(t)dt+\int_{E\setminus E_1}y_2(t)dt.$$

即引理对 $l=2$ 成立。

再设(17)对于 $l-1$ 成立，往证明它对于 l 也成立。

显然，成立着

$$\sum_{j=1}^l\int_E\lambda_j(t)y_j(t)dt=\int_E\sum_{j=1}^{l-1}\lambda_j(t)y_j(t)dt$$

$$+\int_E\lambda_l(t)y_l(t)dt.$$

设 $E'=\{t|\lambda_l(t)=1, t\in E\}$，那末当 $t\in E'$ 时 $\lambda_j(t)=0$, $j=1, \cdots,$

$l-1$. 上式右端等于

$$\int_{E \setminus E'} \sum_{j=1}^{l-1} \lambda_j(t) y_j(t) dt + \int_{E \setminus E'} \lambda_l(t) y_l(t) dt$$

$$+ \int_{E'} y_l(t) dt.$$

上式第一、二项可写为

$$\int_{E \setminus E'} (1-\lambda_l(t)) \sum_{j=1}^{l-1} \frac{\lambda_j(t)}{1-\lambda_l(t)} y_j(t) dt + \int_{E \setminus E'} \lambda_l(t) y_l(t) dt,$$

根据已证，(17)对 $l=2$ 成立，所以上式等于

$$\int_{E'_1} \sum_{j=1}^{l-1} \frac{\lambda_j(t)}{1-\lambda_l(t)} y_j(t) dt + \int_{E'_l} y_l(t) dt,$$

这里 $E'_1 \cup E'_l = E \setminus E'$, $E'_1 \cap E'_l = \varnothing$. 对于上式第一项应用 (17) 对 $l-1$ 成立，得到：存在 $E_j \subset E'_1 (j=1, \cdots, l-1)$，使得

$$E_1 \cup E_2 \cup \cdots \cup E_{l-1} = E'_1,$$

并且 $E_j \cap E_k = \varnothing (j \neq k)$，从而上式等于

$$\sum_{j=1}^{l-1} \int_{E_j} y_j(t) dt + \int_{E'_l} y_l(t) dt.$$

置 $E_l = E'_l \cup E'$，那末

$$E_1 \cup E_2 \cup \cdots \cup E_{l-1} \cup E_l = E,$$

$$E_j \cap E_k = \varnothing \qquad (j \neq k),$$

并且
$$\sum_{j=1}^{l} \int_E \lambda_j(t) y_j(t) dt = \sum_{j=1}^{l} \int_{E_j} y_j(t) dt. \quad \blacksquare$$

定理 8 设 U 是 $\mathbb{R}^m$ 中的有界闭集，那末系统(1)的等时区域 $R(t)$ 是有界凸闭集.

证 设 U_{ad} 表示几乎处处取值于 U 的可积函数的全体；V_{ad} 表示几乎处处取值于 $\operatorname{co} U$ 的可积函数的全体. 显然成立着

$$U_{ad} \subset V_{ad}.$$

记
$$S(t) = \{x \mid x = x(t, v),\ v(\cdot) \in V_{ad}\}.$$

根据定理 2，因为 $\mathrm{co}\,U$ 是有界凸闭集，所以 $S(t)$ 是有界凸闭集.

显然成立着

$$R(t)\subset S(t).$$

现在证明 $R(t)=S(t)$，即如果 $x\in S(t)$，那末 $x\in R(t)$.

事实上，设 $x\in S(t)$，即存在 $v(\cdot)\in V_{ad}$ 使得

$$x=x(t,\ v)=\Phi(t,\ t_0)x_0+\int_{t_0}^{t}\Phi(t,\ s)B(s)v(s)ds.$$

因为当 $s\in[t_0,\ t]$ 时

$$v(s)\in\mathrm{co}\,U,$$

根据定理 2-6，存在

$$u_j(s)\in U\qquad(j=0,\ 1,\ \cdots,\ m),$$

$$\alpha_j(s)\geqslant 0\qquad(j=0,\ 1,\ \cdots,\ m)$$

满足

$$v(s)=\alpha_0(s)u_0(s)+\alpha_1(s)u_1(s)+\cdots+\alpha_m(s)u_m(s),\qquad(*)$$

$$1=\alpha_0(s)+\alpha_1(s)+\cdots+\alpha_m(s).$$

记

$$U_0=\{\alpha=(\alpha_0,\ \alpha_1,\ \cdots,\ \alpha_m)\mid 0\leqslant\alpha_j,\ \alpha_0+\alpha_1+\cdots+\alpha_m=1\},$$

$$Q=U_0\times\underbrace{U\times\cdots\times U}_{m+1}=U_0\times U^{m+1}.$$

那末 Q 是 $\mathbb{R}^{(m+1)^2}$ 中的有界闭集，不依赖于 s，从而关于包含关系是上半连续的. 而由等式

$$f(\alpha,\ u_0,\ u_1,\ \cdots,\ u_m)=\sum_{j=0}^{m}\alpha_j u_j$$

定义的函数 $f:Q\to\mathbb{R}^m$ 是连续的. 并且

$$f(Q)=\mathrm{co}\,U.$$

而 $(*)$ 式表明 $v(s)\in f(Q)$.

因为 $v(\cdot)$ 是可积的，根据附录 4 的菲利浦夫 (А. Ф. Фили-

пнов)关于可测的隐函数引理，可以取 $\alpha_j(\cdot)$ 和 $u_j(\cdot)$ ($j=0, 1, \cdots, m$) 为可测的，并且满足(*)式.

从而

$$\begin{aligned} x &= x(t, v) \\ &= \Phi(t, t_0)x_0 + \sum_{j=0}^{m}\int_{t_0}^{t}\Phi(t, s)B(s)\alpha_j(s)u_j(s)ds \\ &= \Phi(t, t_0)x_0 + \sum_{j=0}^{m}\int_{t_0}^{t}\alpha_j(s)\Phi(t, s)B(s)u_j(s)ds. \end{aligned}$$

根据定理 7，存在可测集 $E_j\subset[t_0, t]$ ($j=0, 1, \cdots, m$) 使得

$$[t_0, t]=\bigcup_{j=0}^{m}E_j,$$

$$E_j\cap E_k=\varnothing \qquad (j\neq k),$$

$$\sum_{j=0}^{m}\int_{t_0}^{t}\alpha_j(s)\Phi(t, s)B(s)u_j(s)ds = \sum_{j=0}^{m}\int_{E_j}\Phi(t, s)B(s)u_j(s)ds.$$

置 $\quad u(s)=u_j(s)$，当 $s\in E_j$ 时，$j=0, 1, \cdots, m$，

那末 $u(s)\in U$，且 $u(\cdot)$ 在 $[t_0, t]$ 上是可积的，从而

$$u(\cdot)\in U_{ad},$$

并且

$$\sum_{j=0}^{m}\int_{t_0}^{t}\alpha_j(s)\Phi(t, s)B(s)u_j(s)ds = \int_{t_0}^{t}\Phi(t, s)B(s)u(s)ds.$$

所以

$$x=\Phi(t, t_0)x_0+\int_{t_0}^{t}\Phi(t, s)B(s)u(s)ds,$$

$$u(\cdot)\in U_{ad}.$$

即 $x\in R(t)$. 因此

$$R(t)=S(t)$$

是有界凸闭集. ∎

注 2 如果用 U^e 表示 $\mathrm{co}\,U$ 的端点的全体，那末根据引理 3,

$\mathrm{co}U^{e}=\mathrm{co}U$. 显然 $U^{e}\subset U$. 设 U_{ad}^{e} 表示几乎处处取值于 U^{e} 的可积函数的全体，记

$$R^{e}(t)=\{x \mid x=x(t,\ u),\ u(\cdot)\in U_{ad}^{e}\}.$$

根据定理 8 和 $\mathrm{co}\,U^{e}=\mathrm{co}\,U$，我们有下述结论

$$R^{e}(t)=R(t)=S(t).$$

由此，证得

定理 9 如果 x 是系统(1) 用容许控制在时刻 t 达到的点，那末存在仅取值于 U^{e} 的容许控制，它也将系统 (1) 于时刻 t 达到 x. ∎

在定理 8 和定理 9 的证明中，李雅普诺夫定理和菲利浦夫关于可测隐函数引理起着重要作用．拉萨尔是第一个用向量值测度研究控制问题的学者．菲利浦夫在研究最优控制的存在性定理时，证明了他关于可测函数的隐函数引理．在定理 8 的证明中，我们利用菲利浦夫引理以保证 $\alpha_j(\cdot)$ 和 $u_j(\cdot)$ 是可积的(如果它们存在的话)．

如果仅讨论 $\bar{R}(t)$ 的凸性，则只要利用定理 I-1-3.

习　　题

1 试决定系统

$$\frac{dx(t)}{dt}=y(t),\qquad \frac{dy(t)}{dt}=u(t),$$

$$x(0)=y(0)=0,$$

$$u(\cdot):[0,\ 4]\to\{-1,\ 0,\ 1\}\ \text{是可积的},$$

的等时区域.

2 设 $\{\varphi_k(\cdot)\}$ 是 $L^2([0,\ 1];\mathbb{R})$ 上的序列，它弱收敛于

$$\varphi(\cdot)\in L^2([0,\ 1];\mathbb{R}).$$

试考察

$$\int_0^1|\varphi_k(t)|^2dt,$$

当 $k\to\infty$ 时的性状。它是否有极限？它是否有界？并说明理由。

3 设 $L^2([0,1];\mathbb{R})$ 上的序列 $\{\varphi_k(\cdot)\}$ 弱收敛于 $\varphi(\cdot)$，并且

$$\lim_{k\to\infty}\int_0^1\{\varphi_k(t)\}^2dt=\int_0^1\{\varphi(t)\}^2dt,$$

试证

$$\lim_{k\to\infty}\int_0^1\{\varphi_k(t)-\varphi(t)\}^2dt=0.$$

4 设 $b(\cdot,\cdot):[t_0,t_1]\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的，$A(\cdot):[t_0,t_1]\to\mathbb{R}^{n\times n}$ 是可积的，$U\subset\mathbb{R}^m$ 是有界凸闭集，试证明：系统

$$\frac{dx(t)}{dt}=A(t)x(t)+b(t,u(t)),$$

$$x(t_0)=x_0,\qquad u(\cdot)\in U_{ad}$$

的等时区域

$$R(t)=\{x\,|\,x=x(t,u),\quad u(\cdot)\in U_{ad}\}$$

的闭包 $\bar{R}(t)$ 是凸集.

5 设 $\qquad\Sigma=\{x(\cdot)\,|\,x(\cdot)=x(\cdot,u),\quad u(\cdot)\in U_{ad}\}$,

试证明 Σ 在 $C([t_0,t_1];\mathbb{R}^n)$ 中的闭包是凸的.

6 对于系统

$$\frac{dx}{dt}=u(t),\qquad x(0)=0,$$

$$U=\{0,1\},$$

$$u(\cdot)\in U_{ad},$$

试问 Σ 是否是 $C([0,1];\mathbb{R})$ 中的凸集？

§4 最优开关控制器的设计原理

设 $A(\cdot):[t_0,t_1]\to\mathbb{R}^{n\times n}$ 和 $B(\cdot):[t_0,t_1]\to\mathbb{R}^{n\times m}$ 是可积的.控制区域 U 是 $\mathbb{R}^m$ 中的有界闭集，初始状态 x_0 是 $\mathbb{R}^n$ 中给定的点.考虑线性系统

$$\frac{dx(t)}{dt}=A(t)x(t)+B(t)u(t),$$

$$x(t_0)=x_0,\tag{1}$$

$$u(\cdot)\in U_{ad},$$

这里 U_{ad} 表示取值于 U 的可积函数的全体.

再设在 $\mathbb{R}^n$ 中存在一个运动的目标，它的运动 $z(\cdot):[t_0, t_1]\to\mathbb{R}^n$ 是给定的连续函数.

在控制问题的讨论中，首先遇到的问题是：是否存在容许控制 $u(\cdot)$，使得对于某个 $\tilde{t}\in[t_0, t_1]$，成立着

$$x(\tilde{t}, u)=z(\tilde{t}); \tag{2}$$

即系统 (1) 能否击中运动的目标 $z(\cdot)$？这是一个有约束的能控性问题. 它相当于问：是否存在 $\tilde{t}\in[t_0, t_1]$，使得

$$z(\tilde{t})\in R(\tilde{t}). \tag{3}$$

本书不讨论该问题，而始终假定存在 $\tilde{t}$ 使(3)成立.

我们要讨论时间最优控制问题：如果存在 $\tilde{t}\in[t_0, t_1]$ 使得(3)成立，讨论系统(1)最快击中目标 $z(\cdot)$ 的问题，即求

$$t^*=\inf\{\tilde{t}\,|\,z(\tilde{t})\in R(\tilde{t})\}. \tag{4}$$

在上述问题中，首先是 t^* 能否达到，即是否存在容许控制 $u^*(\cdot)$，使得成立着

$$x(t^*, u)=z(t^*)$$

或者关系
$$z(t^*)\in R(t^*)$$
是否成立？

第二个问题：如果 t^* 能够达到，那末 t^* 和最优控制 $u^*(\cdot)$ 满足怎样的条件？

在以后的讨论中，我们假设 U 是 $\mathbb{R}^m$ 中的有界闭集. 如果读者对§3的定理5～9不很清楚，不妨认为 U 是 $\mathbb{R}^m$ 中的有界凸闭集，甚至认为 U 是 $\mathbb{R}^m$ 中的有界凸多面体.

定理 1 如果存在 $\tilde{t}$ 使(3)成立，那末时间最优控制是存在的.

证 根据 t^* 的定义，存在 $\{t_k\}$ 满足

$$t_1\geqslant t_2\geqslant\cdots\geqslant t_k\geqslant\cdots,$$

$$\lim_{k\to\infty} t_k=t^*,$$

$$z(t_k)\in R(t_k).$$

这时存在 $u_k(\cdot)\in U_{ad}$，使得

$$z(t_k)=x(t_k,\ u_k).$$

但是 $x(t_k,\ u_k)=x(t^*,\ u_k)+\{x(t_k,\ u_k)-x(t^*,\ u_k)\}$，

而上式右端的第二项等于

$$\int_{t^*}^{t_k}\Phi(t_k,\ s)B(s)u_k(s)ds$$

$$+\int_{t_0}^{t^*}\{\Phi(t_k,\ s)-\Phi(t^*,\ s)\}B(s)u_k(s)ds$$

$$+\{\Phi(t_k,\ t_0)-\Phi(t^*,\ t_0)\}x_0.$$

它当 $k\to\infty$ 时趋于 0，所以

$$\lim_{k\to\infty}x(t_k,\ u_k)=\lim_{k\to\infty}x(t^*,\ u_k).$$

从而
$$z(t^*)=\lim_{k\to\infty}z(t_k)=\lim_{k\to\infty}x(t^*,\ u_k).$$

但是 $x(t^*,\ u_k)\in R(t^*)$，并且 $R(t^*)$ 是闭集，所以

$$z(t^*)\in R(t^*).$$ ∎

定理 1 称为时间最优控制的存在定理.

定理 2 设最优时间为 t^*，那末 $z(t^*)$ 是 $R(t^*)$ 的边界点.

证 因为已证 $z(t^*)\in R(t^*)$，所以只要证明 $z(t^*)$ 不是 $R(t^*)$ 的内点.

不妨假设 $z(t)\equiv 0$. 事实上，若置

$$\tilde{R}(t)=R(t)-z(t),$$

那末 $\tilde{R}(t)$ 在 $t\geqslant 0$ 上是连续变动的，且对于 $t\geqslant 0$，$\tilde{R}(t)$ 是凸闭集. 并且，$z(t)\in R(t)$ 等价于 $0\in\tilde{R}(t)$. 因此，我们下面认为 $z(t)\equiv 0$，证明 0 是 $R(t^*)$ 的边界点.

如果 0 是 $R(t^*)$ 的内点，那末存在 $\delta>0$，使得

$$\Omega=\{x\,|\,\|x\|<\delta\}\subset R(t^*).$$

由于 $R(\cdot)$ 的连续性，存在 $\eta>0$，使得当 $t^*-\eta<t<t^*+\eta$ 时

$$d(R(t),\ R(t^*))<\delta/2. \tag{5}$$

现在证明：当 $t^*-\eta<t<t^*$ 时，成立着

$$0\in R(t).$$

否则，存在 $\tilde{t}\in(t^*-\eta,\ t^*)$，使得

$$0\bar{\in} R(\tilde{t}). \tag{6}$$

由于 $R(\tilde{t})$ 是凸闭集，所以存在超平面把 $R(\tilde{t})$ 和 0 分在两侧。从而在 $\Omega=\{x|\|x\|\leqslant\delta\}$ 中存在点 x_1，使得

$$d(x_1,\ R(\tilde{t}))\geqslant\delta,$$

$$\|x_1\|\leqslant\delta.$$

因为 $x_1\in\Omega\subset R(t^*)$，所以

$$d(R(t^*),\ R(\tilde{t}))\geqslant\frac{1}{2}d(x_1,\ R(\tilde{t}))\geqslant\frac{1}{2}\delta.$$

它与(5)矛盾。

因此，如果 0 是 $R(t^*)$ 的内点，那末

$$0\in R(t)$$

对于 $t\in(t^*-\eta,\ t^*)$ 成立。它与 t^* 的最优性矛盾。因此，0 不是 $R(t^*)$ 的内点，而是 $R(t^*)$ 的边界点。 ▌

定理 3 最优时间 t^* 是非线性方程

$$F(t)=0 \tag{7}$$

的最小根，这里 $F(\cdot)$ 的定义如下：

$$F(t)=\inf_{\|\lambda\|=1}\left\{\langle\lambda,\ \Phi(t,\ t_0)x_0-z(t)\rangle\right.$$
$$\left.+\int_{t_0}^{t}\max_{u\in U}\langle\lambda,\ \Phi(t,\ s)B(s)u\rangle ds\right\}. \tag{8}$$

并且，如果 λ_0 适合

$$\|\lambda_0\|=1, \tag{9}$$

和

$$\langle\lambda_0, \Phi(t^*, t_0)x_0 - z(t^*)\rangle$$

$$+\int_{t_0}^{t^*}\max_{u\in U}\langle\lambda_0, \Phi(t^*, s)B(s)u\rangle ds = 0, \tag{10}$$

那末最优控制 $u^*(\cdot)$ 使得最大值条件

$$\max_{u\in U}\langle\lambda_0, \Phi(t^*, t)B(t)u\rangle = \langle\lambda_0, \Phi(t^*, t)B(t)u^*(t)\rangle \tag{11}$$

在 $[t_0, t^*]$ 上几乎处处成立.

证 当 $t_0 \leqslant t < t^*$ 时,

$$z(t) \bar{\in} R(t).$$

因为 $R(t)$ 是有界凸闭集, 所以存在 $\lambda_1 \in \mathbb{R}^n$ 和实数 c, 使得 $\|\lambda_1\| = 1$, 并且

$$\langle\lambda_1, x(t, u)\rangle \leqslant c < \langle\lambda_1, z(t)\rangle$$

对于 $u(\cdot) \in U_{ad}$ 成立. 从而成立着

$$\left\langle\lambda_1, \Phi(t, t_0)x_0 + \int_{t_0}^{t}\Phi(t, s)B(s)u(s)ds\right\rangle$$

$$\leqslant c < \langle\lambda_1, z(t)\rangle.$$

所以, 当 $u(\cdot) \in U_{ad}$ 时

$$\langle\lambda_1, \Phi(t, t_0)x_0 - z(t)\rangle + \int_{t_0}^{t}\langle\lambda_1, \Phi(t, s)B(s)u(s)\rangle ds$$

$$\leqslant c - \langle\lambda_1, z(t)\rangle < 0. \tag{12}$$

现在证明: 成立着等式

$$\sup_{u(\cdot)\in U_{ad}}\int_{t_0}^{t}\langle\lambda_1, \Phi(t, s)B(s)u(s)\rangle ds$$

$$= \int_{t_0}^{t}\max_{u\in U}\langle\lambda_1, \Phi(t, s)B(s)u\rangle ds. \tag{13}$$

首先, 当 $B(\cdot)$ 连续时, 由 U 是有界闭集, 存在 $\tilde{u}(s)$ 使得

$$\langle\lambda_1, \Phi(t, s)B(s)\tilde{u}(s)\rangle = \max_{u\in U}\langle\lambda_1, \Phi(t, s)B(s)u\rangle.$$

又因为 $\langle\lambda_1, \Phi(t, \cdot)B(\cdot)u(\cdot)\rangle: [t_0, t]\times U \to \mathbb{R}$ 是连续的, 上式右端是可测的, 从而根据附录中的菲利浦夫引理, 可以取 $\tilde{u}(\cdot)$ 为可

测的，使得上式在 $[t_0, t]$ 上几乎处处成立，从而存在 $\tilde{u}(\cdot)\in U_{ad}$，使得

$$\int_{t_0}^{t}\langle\lambda_1, \Phi(t, s)B(s)\tilde{u}(s)\rangle ds = \int_{t_0}^{t}\max_{u\in U}\langle\lambda_1, \Phi(t, s)B(s)u\rangle ds,$$

从而当 $B(\cdot)$ 连续时(13)成立。

其次，存在 C，使得

$$\|\Phi(t, s)\|\leqslant C, \ \|u\|\leqslant C \quad \forall u\in U.$$

对于 $\varepsilon>0$ 和 $B(\cdot)$ 可积的情形，存在 $B_\varepsilon(\cdot)\in C([t_0, t]; \mathbb{R}^{m\times n})$，使得

$$\int_{t_0}^{t}\|B(s)-B_\varepsilon(s)\|ds\leqslant\frac{\varepsilon}{\|\lambda_1\|C^2}. \tag{14}$$

由于

$$\begin{aligned}\langle\lambda_1, \Phi(t, s)B(s)u\rangle &\leqslant\max_{u\in U}\langle\lambda_1, \Phi(t, s)B_\varepsilon(s)u\rangle\\ &\quad+\max_{u\in U}\langle\lambda_1, \Phi(t, s)[B(s)-B_\varepsilon(s)]u\rangle\\ &\leqslant C^2\|\lambda_1\|\|B(s)-B_\varepsilon(s)\|\\ &\quad+\max_{u\in U}\langle\lambda_1, \Phi(t, s)B_\varepsilon(s)u\rangle.\end{aligned}$$

从而

$$\begin{aligned}&\max_{u\in U}\langle\lambda_1, \Phi(t, s)B(s)u\rangle\\ &\quad\leqslant\|B(s)-B_\varepsilon(s)\|\cdot\|\lambda_1\|C^2\\ &\qquad+\max_{u\in U}\langle\lambda_1, \Phi(t, s)B_\varepsilon(s)u\rangle.\end{aligned}$$

由此得到

$$\begin{aligned}&\int_{t_0}^{t}\max_{u\in U}\langle\lambda_1, \Phi(t, s)B(s)u\rangle ds\\ &\quad\leqslant\varepsilon+\int_{t_0}^{t}\max_{u\in U}\langle\lambda_1, \Phi(t, s)B_\varepsilon(s)u\rangle ds.\end{aligned}$$

从而存在 $\tilde{u}_\varepsilon(\cdot)\in U_{ad}$，使得上式右端等于

$$\varepsilon+\int_{t_0}^{t}\langle\lambda_1, \Phi(t, s)B_\varepsilon(s)\tilde{u}_\varepsilon(s)\rangle ds.$$

但是

$$\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)B_\varepsilon(s)\tilde{u}_\varepsilon(s)\rangle ds$$

$$=\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)B(s)\tilde{u}_\varepsilon(s)\rangle ds$$

$$+\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)[B_\varepsilon(s)-B(s)]\tilde{u}_\varepsilon(s)\rangle ds$$

$$\leqslant\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)B(s)\tilde{u}_\varepsilon(s)\rangle ds+\varepsilon.$$

从而

$$\int_{t_0}^{t}\max_{u\in U}\langle\lambda_1,\ \Phi(t,\ s)B(s)u\rangle ds$$

$$\leqslant 2\varepsilon+\sup_{u(\cdot)\in U_{ad}}\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)B(s)u(s)\rangle ds.$$

由于 $\varepsilon>0$ 的任意性，由上式得到

$$\int_{t_0}^{t}\max_{u\in U}\langle\lambda_1,\ \Phi(t,\ s)B(s)u\rangle ds$$

$$\leqslant\sup_{u(\cdot)\in U_{ad}}\int_{t_0}^{t}\langle\lambda_1,\ \Phi(t,\ s)B(s)u(s)\rangle ds.$$

因此，等式(13)对于可积的 $B(\cdot)$ 成立。

根据(12)和(13)，得到

$$\langle\lambda_1,\ \Phi(t,\ t_0)x_0-z(t)\rangle+\int_{t_0}^{t}\max_{u\in U}\langle\lambda_1,\ \Phi(t,\ s)B(s)u\rangle ds$$

$$\leqslant c-\langle\lambda_1,\ z(t)\rangle<0.$$

根据 $F(t)$ 的定义，上式左端不小于 $F(t)$，所以

$$F(t)\leqslant c-\langle\lambda_1,\ z(t)\rangle<0. \tag{15}$$

对于 t^*，因为 $z(t^*)$ 是 $R(t^*)$ 的边界点，所以存在超平面把 $z(t^*)$ 和 $R(t^*)$ 分在两侧. 从而，和上面类似的讨论同样可以得到

$$F(t^*)\leqslant 0. \tag{16}$$

又因为 $z(t^*)\in R(t^*)$，所以存在 $u^*(\cdot)\in U_{ad}$，使得

$$z(t^*)=\Phi(t^*,\ t_0)x_0+\int_{t_0}^{t^*}\Phi(t^*,\ s)B(s)u^*(s)ds,$$

从而，对于 $\lambda\in R^n$，$\|\lambda\|=1$，成立着

$$\langle\lambda, \Phi(t^*, t_0)x_0-z(t^*)\rangle+\int_{t_0}^{t^*}\langle\lambda, \Phi(t^*, t)B(t)u^*(t)\rangle dt=0. \quad (17)$$

但是

$$\max_{u\in U}\langle\lambda, \Phi(t^*, t)B(t)u\rangle\geqslant\langle\lambda, \Phi(t^*, t)B(t)u^*(t)\rangle$$

在 $t_0\leqslant t\leqslant t^*$ 上成立．从而由上式和(17)，得到

$$\langle\lambda, \Phi(t^*, t_0)x_0-z(t^*)\rangle$$
$$+\int_{t_0}^{t^*}\max_{u\in U}\langle\lambda, \Phi(t^*, t)B(t)u\rangle dt\geqslant 0.$$

上式关于 $\|\lambda\|=1$ 取下界，就得到

$$F(t^*)\geqslant 0.$$

它与(16)相结合，推出

$$F(t^*)=0.$$

因此，根据上式和(15)得到 t^* 是

$$F(t)=0$$

的最小根．

如果 λ_0 满足(10)，并且(17)对于 λ_0 是成立的，那末

$$\int_{t_0}^{t^*}\max_{u\in U}\langle\lambda_0, \Phi(t^*, t)B(t)u\rangle dt \quad (18)$$
$$=\int_{t_0}^{t^*}\langle\lambda_0, \Phi(t^*, t)B(t)u^*(t)\rangle dt.$$

又显然成立着不等式

$$\max_{u\in U}\langle\lambda_0, \Phi(t^*, t)B(t)u\rangle\geqslant\langle\lambda_0, \Phi(t^*, t)B(t)u^*(t)\rangle.$$

所以由(18)推出：等式(11)在 $[t_0, t^*]$ 上几乎处处成立．

注 1　如果记

$$\psi(t)=\Phi^T(t^*, t)\lambda_0,$$

那末

$$\frac{d\psi(t)}{dt}\equiv-A^T(t)\psi(t), \quad (19)$$

$$\psi(t^*)=\lambda_0, \qquad \|\lambda_0\|=1,$$

而最大值条件(11)成为

$$\max_{u\in U}\langle\psi(t),\ B(t)u\rangle=\langle\psi(t),\ B(t)u^*(t)\rangle.$$

方程(19)是(1)的共轭方程.

因此，我们证得了下述结论.

最大值原理 设 $u^*(\cdot)$ 是系统(1)击中目标 $z(\cdot)$ 的时间最优控制，那末存在(1)的共轭方程(19)的非平凡解 $\psi(\cdot)$，使得等式

$$\max_{u\in U}\langle\psi(t),\ B(t)u\rangle=\langle\psi(t),\ B(t)u^*(t)\rangle$$

在 $[t_0,\ t^*]$ 上几乎处处成立.

例 1 试求出把系统

$$\frac{d^2x}{dt^2}=u(t), \qquad |u(t)|\leqslant 1, \tag{20}$$

从状态 $(-H,\ 0)$ 最快地转移到状态 $(0,\ 0)$ 的最优控制.

本书绪论中运用矩量问题方法讨论过该问题，现在用最大值原理来讨论它.

解 (20)的状态方程组为

$$\begin{aligned}\frac{dx}{dt}&=y,\\ \frac{dy}{dt}&=u(t),\end{aligned} \tag{21}$$

这时

$$A=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \quad B=\begin{pmatrix}0\\ 1\end{pmatrix}$$

$$U=[-1,\ 1].$$

(21)的共轭方程组为

$$\frac{d\varphi}{dt}=0, \qquad \frac{d\psi}{dt}=-\varphi, \tag{22}$$

所以

$$\varphi=c_1, \qquad \psi=c_2-c_1t,$$

其中 c_1, c_2 为实常数.

最大值原理告诉人们,存在不全为 0 的常数 c_1, c_2, 使得

$$\max_{|u|\leq 1}\{(c_2-c_1t)u\}=(c_2-c_1t)u^*(t),$$

所以

$$u^*(t)=\operatorname{sgn}(c_2-c_1t).$$

由上式知道, 除满足 $c_2-c_1t=0$ 的 t 之外,最优控制 $u^*(\cdot)$ 只能取值 1 或 -1. 又因为 c_2-c_1t 最多有一个零点, 所以最优控制最多改变一次符号.

如果 $u^*(t)=1$, 那末系统成为

$$\frac{dx}{dt}=y,\quad \frac{dy}{dt}=1,$$

从而

$$2ydy=2dx,$$

$$y^2=2(x+c), \tag{23}$$

其中 c 为任意常数. 如果最优轨线最后到达状态 (0, 0) 时的控制作用为 1, 那末它必是抛物线

$$y^2=2x$$

上的一部分, 即弧 $\widehat{OA}$ 上的一段.

同样,如果最优轨线最后到达状态(0, 0)时的控制作用为 -1, 那末它必是抛物线

$$y^2=-2x \tag{24}$$

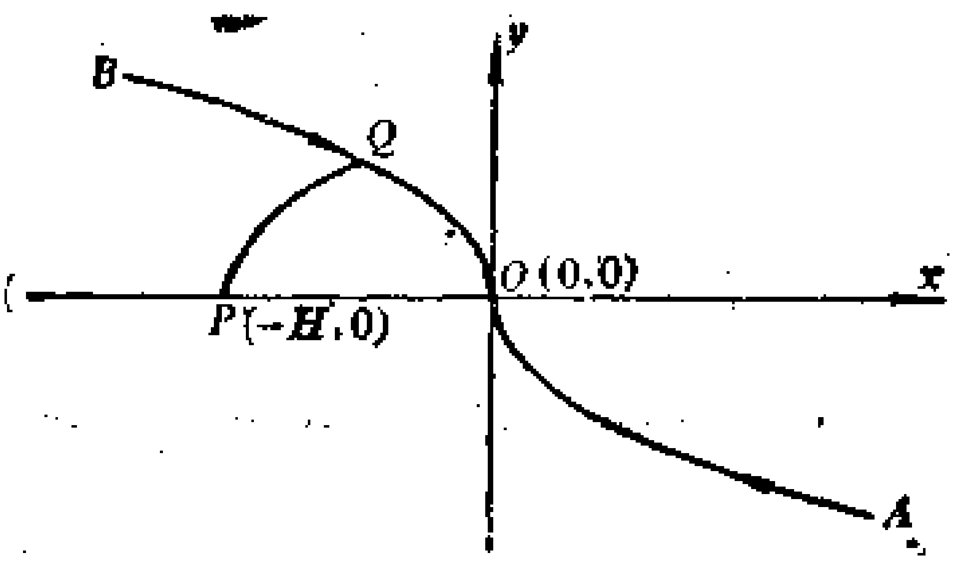


图 2.3

上的一部分，即弧 $\widehat{OB}$ 上的一段.

现在具体分析从状态$(-H,\ 0)$到状态$(0,\ 0)$的最优轨线. 过$(-H,\ 0)$的轨线，一是(23)中的抛物线

$$y^2=2(x+H); \tag{25}$$

另一是系统
$$\frac{dx}{dt}=y,\quad \frac{dy}{dt}=-1$$

过$(-H,\ 0)$的轨线

$$y^2=-2(x+H). \tag{26}$$

仔细分析知：轨线(26)将远离状态$(0,\ 0)$，所以最优轨线段应取(25)的弧段 $\widehat{PQ}$，Q 是抛物线 (25) 和 (24) 的交点 $(-H/2,\ \sqrt{H})$. 然后，最优轨线沿弧 $\widehat{QO}$ 到达状态$(0,\ 0)$. 这样，最优控制 $u^*(\cdot)$ 是

$$u^*(t)=\begin{cases}1, & \text{当 } 0\leqslant t<\sqrt{H} \text{ 时},\\ -1, & \text{当 } \sqrt{H}\leqslant t<2\sqrt{H} \text{ 时},\end{cases}$$

而最优时间 t^* 是

$$t^*=2\sqrt{H}.$$

这正是我们在绪论中得到的结论.

如果置

$$v(x,\ y)=\begin{cases}1, & (x,\ y)\text{在弧 } \widehat{AOB} \text{ 的下方及弧 } \widehat{OA} \text{ 上},\\ -1, & (x,\ y)\text{在弧 } \widehat{AOB} \text{ 的上方及弧 } \widehat{OB} \text{ 上},\end{cases} \tag{27}$$

那末系统

$$\frac{dx}{dt}=y,\quad \frac{dy}{dt}=v(x,\ y) \tag{28}$$

的轨线都是系统(20)的最快到达状态$(0,\ 0)$的最优轨线

(27)的 $v(x,\ y)$称为系统(20)的最优控制的综合函数，(28)是它的综合方程. (28)是状态$(x,\ y)$的非线性反馈.

也就是说，我们由开环最优控制的讨论，得到了闭环最优控制。

下面深入讨论定常线性系统

$$\frac{dx}{dt}=Ax+Bu(t),\quad u(t)\in U,\qquad x(0)=x_0, \tag{29}$$

这里 $A\in\mathbb{R}^{n\times n}$，$B\in\mathbb{R}^{n\times m}$，$U$ 是 $\mathbb{R}^m$ 中的有界凸多面体。并且，假设 U 和 A、B 满足下面的条件：对于平行 U 的某一条棱的向量 w，n 个向量

$$Bw,\ ABw,\ \cdots,\ A^{n-1}Bw \tag{30}$$

是线性无关的。这里，U 的棱是指 U 的边界面的交线。如果条件(30)成立，称 U 与 A、B 处于最广位置。

当 $m=1$，$U=[-1,1]$，$B\in\mathbb{R}^n$ 时，条件(30)成为

$$B,\ AB,\ \cdots,\ A^{n-1}B$$

是线性无关的，它是系统(29)的完全能控性条件。关于线性系统的完全能控性定义和条件，可以参阅《计算机应用中的控制理论》[7]，这里不详述了。

定理 4 设 U 与 A、B 处于最广位置，$\psi(\cdot)\neq 0$ 是(29)的共轭方程

$$\frac{d\psi}{dt}=-A^T\psi$$

的解，那末在 $[t_0,t_1]$ 上除有限个点外，关系式

$$\langle\psi(t),Bu(t)\rangle=\max_{u\in U}\langle\psi(t),Bu\rangle \tag{31}$$

唯一地决定了 $u(t)$，并且 $u(\cdot)$ 在 $[t_0,t_1]$ 上是逐段常值的，而 $u(t)$ 的取值是 U 的顶点。

证 当 $t\in[t_0,t_1]$ 给定时，$\langle\psi(t),Bu\rangle$ 关于 u 是 U 上的线性函数，它或者仅在 U 的边界上达到最大值，或者是常值的。因此，

$\langle\psi(t), Bu\rangle$或者仅在 U 的某一顶点达到最大值，或者在 U 的某条棱上取常值.

现在证明：在 $[t_0, t_1]$ 上除有限个 t 外，由(29)决定的 $u(t)$ 是唯一的.

否则，存在无限个 $t^k(k=1, 2, \cdots)$，使得当 $t=t^k$ 时(31)不能唯一的决定 $u(t^k)$. 这时，$\langle\psi(t^k), Bu\rangle$ 必在 U 的某一条棱上取常值，从而存在 w 平行于 U 的棱使得

$$\langle\psi(t^k), Bw\rangle=\langle\psi(t^k), Bu_0\rangle-\langle\psi(t^k), Bu_1\rangle=0,$$

这里 u_0 和 u_1 是 U 的顶点.

由于 U 仅有有限条棱，所以至少应有无限个 $\{t^{k_j}\}$ 使得：对于同一个 w，成立着

$$\langle\psi(t^{k_j}), Bw\rangle=0. \tag{32}$$

因为 $\psi(\cdot)$ 是常系数线性微分方程组的解，所以

$$\langle\psi(t), Bw\rangle$$

是变量 t 的解析函数. 根据解析函数的零点孤立性定理和(32)推知：当 $t\in[t_0, t_1]$ 时

$$\langle\psi(t), Bw\rangle\equiv0. \tag{33}$$

上式两端对 t 求导，得到

$$\langle-A^T\psi(t), Bw\rangle\equiv0,$$

即

$$\langle\psi(t), ABw\rangle\equiv0.$$

逐次对 t 求导，就得到

$$\langle\psi(t), Bw\rangle\equiv0,$$

$$\langle\psi(t), ABw\rangle\equiv0,$$

$$\langle\psi(t), A^2Bw\rangle\equiv0,$$

$$\cdots\cdots$$

$$\langle\psi(t), A^{n-1}Bw\rangle\equiv0.$$

因为向量组(30)是线性无关的，所以由上式得到

$$\psi(t) \equiv 0.$$

它与 $\psi(\cdot) \neq 0$ 相矛盾。

因此，除有限个 t 外，(31)唯一地决定了 $u(t)$。

这有限个 t 和 t_0、t_1 把 $[t_0, t_1]$ 分成有限个小区间。

设 J 是上述小区间中的一个。

现证：当 $t \in J$ 时 $u(t)$ 是常值的. 事实上，设 U 的顶点为 $e_1, e_2, \cdots, e_q$，而

$$M_j = \{t \mid t \in J,\ u(t) = e_j\}.$$

因此，当 $\bar{t} \in M_j$ 时

$$\langle \psi(\bar{t}),\ Be_j \rangle > \langle \psi(\bar{t}),\ Be_k \rangle \qquad (k \neq j).$$

由于 $\langle \psi(\cdot),\ Be_j \rangle$ 和 $\langle \psi(\cdot),\ Be_k \rangle$ 都是 t 的连续函数，所以由上式得到：存在 $\delta > 0$，使得当 $t \in (\bar{t} - \delta,\ \bar{t} + \delta)$ 时

$$\langle \psi(t),\ Be_j \rangle > \langle \psi(t),\ Be_k \rangle \qquad (k \neq j).$$

这就是说，如果 M_j 不是空集，那末它在 J 中是开集. 但是

$$J = M_1 \cup M_2 \cup \cdots \cup M_q,$$

$$M_j \cap M_k = \varnothing \qquad (j \neq k).$$

因为小区间 J 是个连通集，所以上面的 $M_j (j = 1, 2, \cdots, q)$ 有且只有一个是非空的. 也就是说，当 $t \in J$ 时，$u(t)$ 取常值. ▌

该定理表明：假设 U 与 A、B 处于最广位置，那末由关系(31)决定的 $u(\cdot)$ 是逐段常值的. 又根据最大值原理，时间最优控制是满足形状如(31)的关系的. 因此，最优控制 $u^*(\cdot)$ 在 $[t_0, t^*]$ 上必是逐段常值的，并且它的值是有界凸多面体的顶点. 这样，最优控制是从 U 的一个顶点换接到另一顶点的，并且换接次数是有限的. 该结果称为*开关原理*. 这种类型的控制作用称为*开关控制*.

我们已经知道，当 U 与 A、B 处于最广位置时，满足最大值条件的控制作用是开关控制.

有时，对于一般的线性系统(1)，如果控制作用 $u(\cdot)$ 满足最大

值条件

$$\langle\psi(t),\ B(t)u(t)\rangle=\max_{u\in U}\langle\psi(t),\ B(t)u\rangle,$$

$$\psi(t)\neq 0,$$

就称 $u(\cdot)$ 为开关控制. 显然, 这只是广义的开关控制. 因为这时不能肯定 $u(t)$ 仅取值 U 的顶点, 也不能肯定换接次数是有限的.

习　　题

1　对于系统

$$\frac{d^2x}{dt^2}+x=u(t),$$

$$|u(t)|\leqslant 1,$$

讨论把$(x_0,\ \dot{x}_0)$转移到$(0,\ 0)$的时间最优控制.

2　对于系统

$$\frac{d^nx}{dt^n}=u(t),\quad |u(t)|\leqslant 1,$$

讨论时间最优控制问题.

3　设对于 $t\in[t_0,\ t_1]$, $Q(t)\subset\mathbb{R}^n$ 是有界凸闭集, 并且

$$\lim_{\bar{t}\to t} d(Q(t),\ Q(\bar{t}))=0$$

对于 $t\in[t_0,\ t_1]$ 皆成立, 试讨论: 系统

$$\frac{dx}{dt}=A(t)x+B(t)u(t),$$

$$u(t)\in U,$$

$$x(t_0)=x_0$$

最快击中目标集 $Q(\cdot)$ 的时间最优控制问题, 这里 $A(\cdot)$, $B(\cdot)$和 U 适合 §4 的条件.

4　对于二阶系统

$$\frac{d^2x}{dt^2}=u(t),\quad |u(t)|\leqslant 1,$$

试讨论最快击中目标集 $x=0$ 的时间最优控制问题.

5　对于系统

$$\frac{dx}{dt}=y,\quad \frac{dy}{dt}=-x+u(t),$$

$$|u(t)|\leqslant 1,$$

试讨论最快击中目标集

$$x^2+y^2\leqslant 1$$

的时间最优控制问题.

§5 最优开关控制的唯一性

设 $A\in\mathbb{R}^{n\times n}$, $B\in\mathbb{R}^{n\times m}$, U 是 $\mathbb{R}^m$ 中的有界凸多面体, 并且 U 与 A、B 处于最广位置, 即如果 w 平行于 U 的一条棱, 那末

$$Bw,\ ABw,\ \cdots,\ A^{n-1}Bw \tag{1}$$

是线性无关的.

考虑定常线性系统

$$\frac{dx}{dt}=Ax+Bu(t),\quad u(t)\in U, \tag{2}$$

由状态 x_0 最快转移到状态 x_1 的时间最优控制问题.

定理 1 把系统(2)从状态 x_0 最快转移到状态 x_1 的最优控制是唯一的.

证 设 $u_1(\cdot)$ 和 $u_2(\cdot)$ 都是用最优时间 t^* 把系统 (2) 从状态 x_0 转移到状态 x_1 的控制作用, 即

$$x_1=e^{At^*}x_0+\int_0^{t^*}e^{A(t^*-s)}Bu_1(s)ds,$$

$$x_1=e^{At^*}x_0+\int_0^{t^*}e^{A(t^*-s)}Bu_2(s)ds.$$

因此

$$\int_0^{t^*}e^{A(t^*-t)}Bu_1(t)dt=\int_0^{t^*}e^{A(t^*-t)}Bu_2(t)dt. \tag{3}$$

因为 $u_1(\cdot)$是最优控制, 所以存在共轭方程

$$\frac{d\psi}{dt}=-A^T\psi$$

的非平凡解 $\psi(\cdot)$, 使得等式

$$\langle\psi(t),\ Bu_1(t)\rangle=\max_{u\in U}\langle\psi(t),\ Bu\rangle. \tag{4}$$

在$[0,\ t^*]$上除有限个t外成立．设$\psi(t^*)=\psi_0$，那末

$$\psi(t)=e^{A^T(t^*-t)}\psi_0.$$

用ψ_0和(3)的两端做内积，就得到

$$\int_0^{t^*}\langle\psi_0,\ e^{A(t^*-t)}Bu_1(t)\rangle dt=\int_0^{t^*}\langle\psi_0,\ e^{A(t^*-t)}Bu_2(t)\rangle dt,$$

所以

$$\int_0^{t^*}\langle\psi(t),\ Bu_1(t)\rangle dt=\int_0^{t^*}\langle\psi(t),\ Bu_2(t)\rangle dt.$$

根据(4)，由上式得到

$$\int_0^{t^*}\max_{u\in U}\langle\psi(t),Bu\rangle dt=\int_0^{t^*}\langle\psi(t),\ Bu_2(t)\rangle dt.$$

根据
$$\max_{u\in U}\langle\psi(t),\ Bu\rangle\geqslant\langle\psi(t),\ Bu_2(t)\rangle,$$

由(5)得到：等式

$$\max_{u\in U}\langle\psi(t),\ Bu\rangle=\langle\psi(t),\ Bu_2(t)\rangle$$

在$[0,\ t^*]$上几乎处处成立．

根据定理 4-4，上式唯一地决定$u_2(t)$，所以等式

$$u_1(t)=u_2(t)$$

在$[0,\ t^*]$上几乎处处成立． ■

我们称满足最大值条件

$$\langle\psi(t),\ Bu(t)\rangle=\max_{u\in U}\langle\psi(t),\ Bu\rangle \tag{5}$$

的$u(\cdot)$为极值控制，这里$\psi(\cdot)$是共轭方程的非平凡解，而等式除有限个t外成立．

今后，我们规定容许控制为逐段连续的，并且是右连续的．那末，定理 4-4 表明：当$\psi(\cdot)$给定后，极值控制是唯一的，而且等式在$[t_0,\ t_1]$上处处成立．

定理 4-3 则表明：时间最优控制是极值控制．

定理2 设U与A、B处于最广位置，0是U的内点，$u_1(\cdot)$和$u_2(\cdot)$是把系统(2)从状态x_0转移到状态0的极值控制，那末

$$u_1(\cdot)=u_2(\cdot).$$

证 根据假设，存在$\tilde{t}_1$和$\tilde{t}_2$，使得成立着

$$0=e^{A\tilde{t}_1}x_0+\int_0^{\tilde{t}_1}e^{A(\tilde{t}_1-t)}Bu_1(t)dt,$$

$$0=e^{A\tilde{t}_2}x_0+\int_0^{\tilde{t}_2}e^{A(\tilde{t}_2-t)}Bu_2(t)dt.$$

所以

$$-x_0=\int_0^{\tilde{t}_1}e^{-At}Bu_1(t)dt=\int_0^{\tilde{t}_2}e^{-At}Bu_2(t)dt. \tag{6}$$

如果$\tilde{t}_1\geqslant\tilde{t}_2$. 设$\psi(\cdot)$是共轭方程的非平凡解，使得最大值条件

$$\langle\psi(t),\ Bu_1(t)\rangle=\max_{u\in U}\langle\psi(t),\ Bu\rangle$$

在$[0,\ \tilde{t}_1]$上成立. 设$\psi(0)=\psi_0$，那末

$$\psi(t)=e^{-A^Tt}\psi_0.$$

由(6)得到

$$\int_0^{\tilde{t}_1}\langle e^{-A^Tt}\psi_0,\ Bu_1(t)\rangle dt=\int_0^{\tilde{t}_2}\langle e^{-A^Tt}\psi_0,\ Bu_2(t)\rangle dt,$$

即

$$\int_0^{\tilde{t}_1}\max_{u\in U}\langle\psi(t),\ Bu\rangle dt=\int_0^{\tilde{t}_2}\langle\psi(t),\ Bu_2(t)\rangle dt. \tag{7}$$

由于$0\in U$，所以

$$\max_{u\in U}\langle\psi(t),\ Bu\rangle\geqslant 0,$$

所以由(7)得到

$$\int_0^{\tilde{t}_2}\max_{u\in U}\langle\psi(t),\ Bu\rangle dt\leqslant\int_0^{\tilde{t}_2}\langle\psi(t),\ Bu_2(t)\rangle dt.$$

上式右端的被积函数不大于左端的被积函数，所以等式

$$\max_{u\in U}\langle\psi(t),\ Bu\rangle=\langle\psi(t),\ Bu_2(t)\rangle$$

在$[0,\ \tilde{t}_2]$上几乎处处成立. 根据定理4-4,得到

$$u_1(t)=u_2(t) \tag{8}$$

在$[0,\ \tilde{t}_2]$上成立. 同时推出

$$\int_{\tilde{t}_2}^{\tilde{t}_1}\max_{u\in U}\langle\psi(t),\ Bu\rangle dt=0,$$

从而

$$\max_{u\in U}\langle\psi(t),\ Bu\rangle=0$$

在$[\tilde{t}_2,\ \tilde{t}_1]$上成立.

如果$\tilde{t}_1>\tilde{t}_2$, 由上式得到

$$u(t)=0$$

在$[\tilde{t}_2,\ \tilde{t}_1]$上成立, 但0是$U$的内点,它与$u_1(\cdot)$是极值控制矛盾. 所以$\tilde{t}_1=\tilde{t}_2$.

在$[0,\ \tilde{t}_2]$上我们已证(8). ∎

在证明了定理2后,我们可以得到判断最优控制的充分条件. 我们知道,最大值原理是给出最优控制的必要条件,即时间最优控制必是极值控制. 现在, 如果极值控制把系统(2)从状态x_0转移到状态0,根据定理2它是把系统(2)从状态x_0转移到状态0的唯一的极值控制,所以它必是最优控制.

这样, 在定理2的条件下, 为讨论最快击中状态0的最优控制,只要讨论满足要求的极值控制就可以了.

§6 状态变量为线性的系统

前面所讨论的系统,关于状态变量和控制变量都是线性的.自然,由于控制作用是有约束的,所以它对控制变量的线性并不是完全的. 人们自然要问, 前面的结论对于控制变量为非线性的系统是否成立?

考虑系统

$$\frac{dx(t)}{dt}=A(t)x(t)+b(t,\ u(t)), \tag{1}$$
$$x(t_0)=x_0,\quad u(t)\in U.$$

这里假设：

1° $A(\cdot):[t_0,\ t_1]\to\mathbb{R}^{n\times n}$ 是可积的；

2° 对于 $u\in\mathbb{R}^m$，$b(\cdot,\ u):[t_0,\ t_1]\to\mathbb{R}^n$ 是可积的；对于 $t\in[t_0,\ t_1]$，$b(t,\ \cdot):\mathbb{R}^m\to\mathbb{R}^n$ 是连续的；对于 $\mathbb{R}^m$ 中的任一有界集 K，存在 $\mu(\cdot):[t_0,\ t_1]\to\mathbb{R}$ 是可积的，使得：当 $(t,\ u)\in[t_0,t_1]\times K$ 时，成立着不等式

$$\|b(t,\ u)\|\leqslant\mu(t);$$

3° 控制区域 U 是 $\mathbb{R}^m$ 中的给定集．如果有界可积的 $u(\cdot)$：$[t_0,\ t_1]\to\mathbb{R}^m$ 使得 $u(t)\in U$ 在 $[t_0,\ t_1]$ 上几乎处处成立，就称 $u(\cdot)$ 为容许控制，记为

$$u(\cdot)\in U_{ad};$$

在(1)中简记它为 $u(t)\in U$；

4° 初值 $(t_0,\ x_0)\in\mathbb{R}\times\mathbb{R}^n$ 给定．

显然，当 $B(\cdot):[t_0,\ t_1]\to\mathbb{R}^{n\times m}$ 可积，$b(t,\ u)=B(t)u$ 时，(1) 成为前面讨论过的系统．

根据定理 I-1-7，当 $u(\cdot)\in U_{ad}$ 时

$$b(\cdot,\ u(\cdot))$$

在 $[t_0,\ t_1]$ 上是可积的．因此，对于 $u(\cdot)\in U_{ad}$，系统(1)的解可以表示为

$$x(t,\ u)=\Phi(t,\ t_0)x_0+\int_{t_0}^{t}\Phi(t,\ s)b(s,\ u(s))ds, \tag{2}$$

其中 Φ 是齐次线性方程组

$$\frac{dx}{dt}=A(t)x$$

的转移矩阵．

集合

$$R(t)=\{x\,|\,x=x(t,\ u),\ u(\cdot)\in U_{ad}\} \tag{3}$$

称为系统(1)的等时区域.

定理 1 $R(t)$ 和 $\overline{R}(t)$ 都是凸的.

它的证明类似 § 4 的定理 4 和定理 6，读者自行完成.

现在假设 U 是 $\mathbb{R}^m$ 中的有界闭集，记

$$V(t)=\{x\,|\,x=b(t,\ u),\ u\in U\}. \tag{4}$$

因为 U 是有界闭集，$b(t,\ \cdot):\mathbb{R}^m\to\mathbb{R}^n$ 是连续的，所以当 $t\in[t_0,t_1]$ 时，$V(t)$ 是 $\mathbb{R}^n$ 中的闭集.

根据菲利浦夫引理(见附录 4)，如果 $v(\cdot):[t_0,\ t_1]\to\mathbb{R}^n$ 是可积的，并且 $v(t)\in V(t)$ 在 $[t_0,\ t_1]$ 上几乎处处成立，那末存在可积的 $u(\cdot):[t_0,\ t_1]\to\mathbb{R}^m$，使得

$$u(t)\in U$$

和

$$v(t)=b(t,\ u(t)) \tag{5}$$

在 $[t_0,\ t_1]$ 上几乎处处成立．这时记 $v(\cdot)$ 为 $v(\cdot)\in\tilde{V}_{ad}$.

如果 $v(\cdot):[t_0,\ t_1]\to\mathbb{R}^n$ 是可积的，并且 $v(t)\in\mathrm{co}\,V(t)$ 在 $[t_0,t_1]$ 上几乎处处成立，就记它为

$$v(\cdot)\in V_{ad}.$$

对应于(1)，讨论线性系统

$$\frac{d\tilde{x}}{dt}=A(t)\tilde{x}+v(t), \tag{6}$$

$$\tilde{x}(t_0)=x_0.$$

记
$$R_1(t)=\{x\,|\,x=\tilde{x}(t,\ v),\ v(\cdot)\in\tilde{V}_{ad}\},$$
$$S_1(t)=\{x\,|\,x=\tilde{x}(t,\ v),\ v(\cdot)\in V_{ad}\}.$$

利用 § 3 的方法，容易证明

$$R_1(t)=S_1(t).$$

定理 2 成立着等式

$$R(t)=R_1(t)=S_1(t).$$

证 设 $x\in R(t)$，就存在 $u(\cdot)\in U_{ad}$，使得

$$x=x(t,\ u).$$

记 $v(\cdot)=b(\cdot,\ u(\cdot))$，那末 $v(\cdot)\in \tilde{V}_{ad}$，并且

$$\tilde{x}(t,\ v)=x(t,\ u).$$

因此 $$R(t)\subset R_1(t).$$

反之，如果 $v(\cdot)\in \tilde{V}_{ad}$，就存在 $u(\cdot)\in U_{ad}$，使得

$$b(\cdot,\ u(\cdot))=v(\cdot),$$

从而 $$x(t,\ u)=\tilde{x}(t,\ v).$$

所以 $$R_1(t)\subset R(t).$$

因此，成立着

$$R(t)=R_1(t)=S_1(t).$$ ▮

定理 3 设 U 是 $\mathbb{R}^m$ 中的有界闭集，那末 $R(t)$ 是有界凸闭集.

证 和 § 3 一样，可以证明 $S_1(t)$ 是凸闭集. 请读者自行完成. 再由定理 2 就得证定理. ▮

设 $Q(\cdot)$ 满足下述条件：

5° 对于 $t\in[t_0,\ t_1]$，$Q(t)$ 是 $\mathbb{R}^n$ 中的有界凸闭集；

6° $Q(\cdot)$ 关于 $t\in[t_0,\ t_1]$ 是连续的，即成立着

$$\lim_{\bar{t}\to t} d(Q(t),\ Q(\bar{t}))=0.$$

称 $Q(\cdot)$ 是运动着的目标集.

时间最优控制问题 寻求

$$t^*=\inf\{\bar{t}\,|\,x(\bar{t},\ u)\in Q(\bar{t}),\ u(\cdot)\in U_{ad}\}.$$

显然，它等价于寻求

$$\begin{aligned}t^*&=\inf\{\bar{t}\,|\,R(\bar{t})\cap Q(\bar{t})\neq\varnothing\}\\&=\inf\{\bar{t}\,|\,0\in R(\bar{t})-Q(\bar{t})\}.\end{aligned}$$

因为 $R(t)-Q(t)$ 是有界凸闭集，并且 $R(\cdot)-Q(\cdot)$ 关于 t 是连续变动的，所以上述问题可归结为 § 4 讨论的问题，从而有下述定理.

定理 4 如果存在 $\bar{t}$，使得

$$Q(\bar{t})\cap R(\bar{t})\neq\varnothing,$$

那末
$$Q(t^*)\cap R(t^*)\neq\varnothing,$$

并且当 $t_0\leqslant t<t^*$ 时

$$Q(t)\cap R(t)=\varnothing, \tag{7}$$

同时 0 是 $R(t^*)-Q(t^*)$ 的边界点.

定理 5 设 U 是 $\mathbb{R}^m$ 中的有界闭集，那末 t^* 是方程

$$F(t)=0$$

的最小根，这里

$$F(t)=\min_{\|\lambda\|=1}\left\{\max_{x\in Q(t)}\langle\lambda,\ \Phi(t,\ t_0)x_0-x\rangle+\int_{t_0}^{t}\max_{u\in U}\langle\lambda,\ \Phi(t,\ s)b(s,\ u)\rangle ds\right\}. \tag{8}$$

又，如果 λ_0 满足 $\|\lambda_0\|=1$ 及

$$\max_{x\in Q(t^*)}\langle\lambda_0,\ \Phi(t^*,\ t_0)x_0-x\rangle+\int_{t_0}^{t^*}\max_{u\in U}\langle\lambda_0,\ \Phi(t^*,\ t)b(t,\ u)\rangle dt=0,$$

那末最优控制 $u^*(\cdot)$ 满足最大值条件：等式

$$\langle\lambda_0,\ \Phi(t^*,\ t)b(t,\ u^*(t))\rangle=\max_{u\in U}\langle\lambda_0,\ \Phi(t^*,\ t)b(t,\ u)\rangle \tag{9}$$

在 $[t_0,\ t^*]$ 上几乎处处成立，并且对于 $x\in Q(t^*)$，成立着

$$\langle\lambda_0,\ x-x(t^*,\ u^*)\rangle\geqslant 0. \tag{10}$$

它们的证明留给读者完成.

不等式(10)称为横截条件.

相应于系统(1)，还可以考虑它的松弛系统

$$\frac{dx}{dt}=A(t)x+\sum_{j=0}^{n}\alpha_j(t)b(t,\ u_j(t)),\tag{11}$$

这里

$$u_j(\cdot)\in U_{ad}\quad (j=0,\ 1,\ \cdots,\ n);$$

$$\alpha_j(\cdot):[t_0,\ t_1]\to[0,\ 1] \text{ 是可积的;}$$

$$\sum_{j=0}^{n}\alpha_j(t)=1.$$

我们称 $\alpha_j(\cdot)(j=0,\ 1,\ \cdots,\ n)$ 为松弛因子. (11) 的等时区域记为 $R_2(t)$. 显然成立着

$$R(t)\subset R_2(t).$$

容易证明

定理 6 系统(1)的松弛系统(11)的等时区域

$$R_2(t)=\Big\{x\,\Big|\,x=\Phi(t,\ t_0)x_0+\sum_{j=0}^{n}\int_{t_0}^{t}\alpha_j(s)\Phi(t,\ s)b(s,\ u_j(s))ds,$$

$$u_j(\cdot)\in U_{ad},\ \alpha_j(\cdot):[t_0,\ t]\to[0,\ 1]\text{可积},\sum_{j=0}^{n}\alpha_j(s)=1\Big\}$$

与系统(1)的等时区域 $R(t)$ 是相等的，即 $R_2(t)=R(t)$.

它的证明留给读者完成.

习　　题

1　证明本节的定理 1, 3, 4, 5 和 6.

2　设 $R_1(t)$ 和 $R_2(t)$ 都是 $\mathbb{R}^n$ 中的有界凸闭集，并且

$$R_1(t_0)\cap R_2(t_0)=\varnothing,$$

存在 $\bar{t}\in[t_0,\ t_1]$, 使得

$$R_1(\bar{t})\cap R_2(\bar{t})\neq\varnothing,$$

又设 $R_1(\cdot)$ 和 $R_2(\cdot)$ 在$[t_0,\ t_1]$上关于 t 是连续变动的. 试讨论下述最优问题:

$$\inf\{t\,|\,R_1(t)\cap R_2(t)\neq\varnothing\}.$$

§7 最优线性反馈调节器的设计原理

在绪论中，我们曾用配方法讨论最简单的反馈调节器的设计问题．本节讨论它的一般化问题，这是卡尔曼完整解决的，它是最优控制理论中能解到底的最重要的一类问题．

设系统的状态方程为

$$\begin{aligned}&\frac{dx}{dt}=Ax+Bu(t),\\&x(0)=x_0,\end{aligned}\tag{1}$$

这里 $A\in\mathbb{R}^{n\times n}$, $B\in\mathbb{R}^{n\times m}$, $x_0\in\mathbb{R}^n$, $u(\cdot)\in L^2([0,\ \infty);\ \mathbb{R}^m)$.

系统(1)中的 A 可能是不稳定的，即 A 的特征值不全具有负实部．反馈调节器的设计，是要选取 $K\in\mathbb{R}^{m\times n}$，使得 $A-BK$ 是稳定的，并且使系统的性能(品质)较好．在控制工程中，经常以下述的平方积分指标

$$J(u(\cdot))=\int_0^\infty\left\langle\begin{pmatrix}Q & S^T\\ S & R\end{pmatrix}\begin{pmatrix}x(t)\\ u(t)\end{pmatrix},\ \begin{pmatrix}x(t)\\ u(t)\end{pmatrix}\right\rangle dt \tag{2}$$

来衡量系统的优劣，这里

$$\begin{aligned}&Q^T=Q\in\mathbb{R}^{n\times n},\\&S\in\mathbb{R}^{m\times n},\\&R^T=R\in\mathbb{R}^{m\times m},\end{aligned}$$

并且

$$R>0, \tag{3}$$

$$\begin{pmatrix}Q & S^T\\ S & R\end{pmatrix}\geqslant 0. \tag{4}$$

通常 $S=0$.

我们知道，当矩阵

$$(B\,\vdots\, AB\,\vdots\,\cdots\,\vdots\, A^{n-1}B)$$

的秩为 n 时，存在 $K\in\mathbb{R}^{m\times n}$ 使得 $A-BK$ 是稳定的．(参见«计算

机应用中的控制理论»)．在系统(1)中，置

$$u(t)=-Kx(t)+v(t),$$

那末系统的状态方程成为

$$\frac{dx}{dt}=(A-BK)x+Bv(t),$$

而性能指标(2)变为

$$J(v)=\int_0^\infty\left\langle\begin{pmatrix}\tilde{Q} & \tilde{S}^T\\ \tilde{S} & R\end{pmatrix}\begin{pmatrix}x(t)\\ v(t)\end{pmatrix},\ \begin{pmatrix}x(t)\\ v(t)\end{pmatrix}\right\rangle dt,$$

其中 $$\tilde{Q}=Q+K^TRK-S^TK-K^TS,$$

$$\tilde{S}=S-RK,$$

并且 $$\begin{pmatrix}\tilde{Q} & \tilde{S}^T\\ \tilde{S} & R\end{pmatrix}\geqslant 0.$$

因此，我们假定 A 是稳定的，Q, S, R 满足条件(3)和(4)．当 $u(\cdot)\in L^2([0,\ \infty);\ \mathbb{R}^m)$ 时

$$x(t,\ u)=e^{At}x_0+\int_0^t e^{A(t-s)}Bu(s)ds, \tag{5}$$

并且 $x(\cdot,u)\in L^2([0,\ \infty);\ \mathbb{R}^n)$，从而可以计算性能指标 $J(u(\cdot))$

为设计最优反馈调节器问题，我们先讨论下述的二次最优控制问题：寻求 $u^*(\cdot)\in L^2([0,\ \infty);\ \mathbb{R}^m)$，使得不等式

$$J(u^*(\cdot))\leqslant J(u(\cdot))$$

对于 $u(\cdot)\in L^2([0,\ \infty);\ \mathbb{R}^m)$ 成立．

记 $$U_{ad}=L^2([0,\ \infty);\ \mathbb{R}^m).$$

定理 1 设 A 是稳定的，即 A 的特征值全在左半平面，$R>0$，$Q\geqslant 0$，$S=0$，那末线性系统(1)的二次最优控制问题的解是存在唯

置 $$\alpha=\inf_{u(\cdot)\in U_{ad}} J(u(\cdot)),$$

那末 $$\alpha\geqslant 0.$$

设$\{u_k(\cdot)\}$是极小化序列，即

$$u_k(\cdot)\in U_{ad}\qquad (k=1,\ 2,\ \cdots),$$

$$\lim_{k\to\infty} J(u_k(\cdot))=\alpha.$$

记 $$x_k(\cdot)=x(\cdot,\ u_k),$$

那末 $$x\left(\cdot,\ \frac{u_k+u_l}{2}\right)=\frac{1}{2}\{x_k(\cdot)+x_l(\cdot)\}.$$

根据二次型的平行四边形法则，有

$$J(u_k(\cdot))+J(u_l(\cdot))=2J\left(\frac{u_k(\cdot)+u_l(\cdot)}{2}\right)$$

$$+\int_0^\infty 2\left\langle Q\frac{x_k(t)-x_l(t)}{2},\ \frac{x_k(t)-x_l(t)}{2}\right\rangle dt$$

$$+\int_0^\infty 2\left\langle R\frac{u_k(t)-u_l(t)}{2},\ \frac{u_k(t)-u_l(t)}{2}\right\rangle dt$$

上式右端第二项是非负的，所以

$$J(u_k(\cdot))+J(u_l(\cdot))$$

$$\geqslant 2J\left(\frac{(u_k(\cdot)+u_l(\cdot))}{2}\right)$$

$$+2\int_0^\infty\left\langle R\frac{u_k(t)-u_l(t)}{2},\ \frac{u_k(t)-u_l(t)}{2}\right\rangle dt.$$

因为$R>0$，存在$\delta>0$，使得

$$\langle Ru,\ u\rangle\geqslant\delta\|u\|^2,$$

所以 $$J(u_k(\cdot))+J(u_l(\cdot))\geqslant 2\alpha+\frac{\delta}{2}\int_0^\infty\|u_k(t)-u_l(t)\|^2dt.$$

即 $$\int_0^\infty\|u_k(t)-u_l(t)\|^2dt\leqslant\frac{2}{\delta}\{J(u_k(\cdot))+J(u_l(\cdot))-2\alpha\}.$$

当 $k,\ l\to\infty$ 时，上式右端趋于 0，所以 $\{u_k(\cdot)\}$ 是空间 $L^2([0,\ \infty);\ \mathbb{R}^m)$ 中的柯西(A. L. Cauchy)序列，从而存在

$u^*(\cdot)\in L^2([0,\infty);\mathbb{R}^m)$，使得

$$\lim_{k\to\infty}\int_0^\infty\|u_k(t)-u^*(t)\|^2dt=0.$$

又根据(5)，得到

$$x_k(t)-x(t,u^*)=\int_0^t e^{A(t-s)}B\{u_k(s)-u^*(s)\}ds.$$

从而
$$\lim_{k\to\infty}\int_0^\infty\|x_k(t)-x(t,u^*)\|^2dt=0.$$

因此
$$J(u^*(\cdot))=\lim_{k\to\infty}J(u_k(\cdot))=\alpha.$$

所以 $u^*(\cdot)$ 是最优控制．这就证明了最优控制的存在性．再证最优控制的唯一性．如果 $\tilde{u}^*(\cdot)$ 也是最优控制，那末

$$J(u^*(\cdot))=\alpha=J(\tilde{u}^*(\cdot)),$$

$$\begin{aligned}&J(u^*(\cdot))+J(\tilde{u}^*(\cdot))\\&\geqslant 2J\left(\frac{u^*(\cdot)+\tilde{u}^*(\cdot)}{2}\right)\\&\quad+2\int_0^\infty\left\langle R\frac{u^*(t)-\tilde{u}^*(t)}{2},\ \frac{u^*(t)-\tilde{u}^*(t)}{2}\right\rangle dt\\&\geqslant 2\alpha+\frac{\delta}{2}\int_0^\infty\|u^*(t)-\tilde{u}^*(t)\|^2dt,\end{aligned}$$

所以
$$\int_0^\infty\|u^*(t)-\tilde{u}^*(t)\|^2dt\leqslant 0.$$

从而
$$\tilde{u}^*(t)\equiv u^*(t),$$

在$[0,\infty)$上几乎处处成立． ∎

定理 2 设对称阵 P 适合

$$Q+A^TP+PA-PBR^{-1}B^TP=0, \tag{6}$$

并且 $A-BR^{-1}B^TP$ 的特征值全具负实部，那末

$$u^*(t)=-R^{-1}B^TPx^*(t)$$

是最优控制．

证 设 $u(\cdot)\in U_{ad}$，那末 $x(\cdot)\in L^2([0,\infty);\mathbb{R}^n)$．考虑

$$\frac{d}{dt}\langle Px(t),\ x(t)\rangle$$
$$=\langle (A^T P + PA)x(t),\ x(t)\rangle$$
$$+2\langle B^T Px(t),\ u(t)\rangle.$$

两边从 0 到 ∞ 积分，得到

$$-\langle Px_0,\ x_0\rangle = \int_0^\infty \{\langle (A^T P + PA)x(t),\ x(t)\rangle$$
$$+2\langle B^T Px(t),\ u(t)\rangle\}dt.$$

把上式与 $J(u(\cdot))$ 相加，得到

$$J(u(\cdot)) - \langle Px_0,\ x_0\rangle$$
$$= \int_0^\infty \langle (Q + A^T P + PA)x(t),\ x(t)\rangle dt$$
$$+ \int_0^\infty \{2\langle B^T Px(t),\ u(t)\rangle + \langle Ru(t),\ u(t)\rangle\}dt.$$

利用 P 适合 (6)，上式右端第一项等于

$$\int_0^\infty \langle PBR^{-1}B^T Px(t),\ x(t)\rangle dt.$$

所以

$$J(u(\cdot)) - \langle Px_0,\ x_0\rangle$$
$$= \int_0^\infty \langle R(u(t) + R^{-1}B^T Px(t)),$$
$$u(t) + R^{-1}B^T Px(t)\rangle dt \geqslant 0. \tag{7}$$

因此 $$J(u(\cdot)) \geqslant \langle Px_0,\ x_0\rangle.$$

如果 $u^*(\cdot)$ 和 $x^*(\cdot)$ 适合

$$u^*(t) + R^{-1}B^T Px^*(t) = 0, \tag{8}$$

那末 $$\frac{dx^*(t)}{dt} = (A - BR^{-1}B^T P)x^*(t),$$

因为 $A - BR^{-1}B^T P$ 是稳定的，所以 $x^*(\cdot) \in L^2([0,\ \infty);\ \mathbb{R}^n)$，$u^*(\cdot) \in L^2([0,\ \infty);\ \mathbb{R}^m)$．这时，由 (7) 得到

$$J(u^*(\cdot)) - \langle Px_0,\ x_0\rangle = 0.$$

这样，我们证得

$$J(u^*(\cdot))\leqslant J(u(\cdot)).$$ ∎

该定理表明，使性能指标 $J(u(\cdot))$ 取最小的最优控制是状态的线性反馈(8)．这里是推广绪论的配平方的方法来得到结论的．

定理 3 设 $u^*(\cdot)$, $x^*(\cdot)=x(\cdot, u^*)$是线性二次最优控制问题的解，又设

$$\psi(t)=-\int_t^\infty e^{A^T(\sigma-t)}Qx^*(\sigma)d\sigma, \tag{9}$$

那末

$$u^*(t)=R^{-1}B^T\psi(t). \tag{10}$$

证 设 $u(\cdot)\in U_{ad}$, $\varepsilon\in(0, 1)$, 那末 $u^\varepsilon(\cdot)=u^*(\cdot)+\varepsilon u(\cdot)\in U_{ad}$, 从而

$$J(u^*(\cdot))\leqslant J(u^\varepsilon(\cdot)).$$

记 $x^\varepsilon(\cdot)=x(\cdot, u^\varepsilon)=x^*(\cdot)+\varepsilon\delta x(\cdot)$, 而 $\delta x(\cdot)$适合

$$\frac{d\delta x(t)}{dt}=A\delta x(t)+Bu(t),\quad \delta x(0)=0. \tag{11}$$

这样

$$\begin{aligned}J(u^\varepsilon(\cdot))&=J(u^*(\cdot))\\&+2\varepsilon\int_0^\infty\{\langle Qx^*(t), \delta x(t)\rangle+\langle Ru^*(t), u(t)\rangle\}dt\\&+\varepsilon^2\int_0^\infty\{\langle Q\delta x(t), \delta x(t)\rangle+\langle Ru(t), u(t)\rangle\}dt\\&\geqslant J(u^*(\cdot)).\end{aligned}$$

所以

$$2\varepsilon\int_0^\infty\{\langle Qx^*(t), \delta x(t)\rangle+\langle Ru^*(t), u(t)\rangle\}dt+\varepsilon^2\int_0^\infty\{\langle Q\delta x(t), \delta x(t)\rangle+\langle Ru(t), u(t)\rangle\}dt\geqslant0.$$

两端除以 2ε, 并令 $\varepsilon\to0$, 得到

$$\int_0^\infty\{\langle Qx^*(t), \delta x(t)\rangle+\langle Ru^*(t), u(t)\rangle\}dt\geqslant0,$$

但是
$$\delta x(t)=\int_0^t e^{A(t-s)}Bu(s)ds,$$
代入上式得到
$$\int_0^\infty\langle Qx^*(\sigma),\int_0^\sigma e^{A(\sigma-t)}Bu(t)dt\rangle d\sigma$$
$$+\int_0^\infty\langle Ru^*(t),u(t)\rangle dt\geqslant 0.$$
从而
$$\int_0^\infty\langle B^T\int_t^\infty e^{A^T(\sigma-t)}Qx^*(\sigma)d\sigma+Ru^*(t),u(t)\rangle dt\geqslant 0$$
对于 $u(\cdot)\in U_{ad}$ 成立，即
$$\int_0^\infty\langle Ru^*(t)-B^T\psi(t),u(t)\rangle dt\geqslant 0. \tag{12}$$

因为 $u^*(\cdot)\in L^2([0,\infty);\mathbb{R}^m)$，所以 $x^*(\cdot)\in L^2([0,\infty);\mathbb{R}^n)$，$\psi(\cdot)\in L^2([0,\infty);\mathbb{R}^n)$．从而 $Ru^*(\cdot)-B^T\psi(\cdot)\in L^2([0,\infty),\mathbb{R}^n)$．取 $u(\cdot)=B^T\psi(\cdot)-Ru^*(\cdot)$，由上式得到
$$-\int_0^\infty\|Ru^*(t)-B^T\psi(t)\|^2dt\geqslant 0.$$
所以得到等式(10)． ∎

如果置
$$\mathscr{H}(\psi,x,u)=-\frac{1}{2}\{\langle Qx,x\rangle+\langle Ru,u\rangle\}$$
$$+\langle\psi,Ax+Bu\rangle, \tag{13}$$
那末对于给定的 $t>0$，关于 u 的函数
$$\mathscr{H}(\psi(t),x^*(t),u)=-\frac{1}{2}\{\langle Qx^*(t),x^*(t)\rangle+\langle Ru,u\rangle\}$$
$$+\langle\psi(t),Ax^*(t)+Bu\rangle$$
是二次函数，并且
$$\frac{\partial\mathscr{H}(\psi(t),x^*(t),u^*(t))}{\partial u}=-Ru^*(t)+B^T\psi(t)=0,$$
$$\frac{\partial^2\mathscr{H}(\psi(t),x^*(t),u^*(t))}{\partial u^2}=-R<0,$$

所以 $\mathscr{H}(\psi(t),\ x^*(t),\ \cdot)$ 在 $u^*(t)$ 处达到最大值，即有

$$\mathscr{H}(\psi(t),\ x^*(t),\ u^*(t))=\max_{u\in\mathbb{R}^m}\mathscr{H}(\psi(t),\ x^*(t),\ u). \tag{14}$$

因此，我们证得

定理 4 设 $u^*(\cdot)$，$x^*(\cdot)$ 是线性二次最优控制问题的解，$\psi(\cdot)$ 由(9)决定，那末哈密顿(W. R. Hamilton)函数

$$\mathscr{H}(\psi(t),\ x^*(t),\ u)=-\frac{1}{2}\{\langle Qx^*(t),\ x^*(t)\rangle+\langle Ru,\ u\rangle\}$$
$$+\langle\psi(t),\ Ax^*(t)+Bu\rangle$$

在 $u=u^*(t)$ 达到最大值，即等式 (14) 在 $[0,\ +\infty)$ 上几乎处处成立.

习　　题

1　设 $Q^T=Q\in\mathbb{R}^{n\times n}$，试证平行四边形法则

$$\langle Qx,\ x\rangle+\langle Qy,\ y\rangle$$
$$=2\left\{\left\langle Q\frac{x+y}{2},\ \frac{x+y}{2}\right\rangle+\left\langle Q\frac{x-y}{2},\ \frac{x-y}{2}\right\rangle\right\}.$$

2　在区间 $[0,\ T]$ 上考虑线性系统(1)使二次性能指标

$$J(u(\cdot))=\langle Q_0x(T),\ x(T)\rangle+\int_0^T\{\langle Qx(t),\ x(t)\rangle+\langle Ru(t),\ u(t)\rangle\}dt$$

取最小的最优控制问题，这里 $Q_0^T=Q_0\in\mathbb{R}^{n\times n}$，$Q^T=Q\in\mathbb{R}^{n\times n}$，$R^T=R\in\mathbb{R}^{m\times m}$，$Q_0\geqslant0$，$Q\geqslant0$，$R>0$.

1° 证明最优控制的存在唯一性.

2° 证明最优控制适合最大值条件.

3° 设 $P^T(\cdot)=P(\cdot)$ 在 $[0,\ T]$ 上满足黎卡提方程

$$-\frac{dP(t)}{dt}=Q+A^TP(t)+P(t)A-P(t)BR^{-1}B^TP(t),$$
$$P(T)=Q_0,$$

如果 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,\ u)$ 满足关系

$$u(t)=-R^{-1}B^TP(t)x(t),$$

那末 $u(\cdot)$ 是最优控制.

§8 一个线性最优控制问题

对于线性系统

$$\frac{dx}{dt}=A(t)x+b(t,\ v(t)),\qquad x(t_0)=x_0,\ v(t)\in U,\tag{1}$$

假设 $A(\cdot)$, $b(\cdot,\ \cdot)$ 满足 §6 所述条件 1° 和 2°, U 是 $\mathbb{R}^m$ 中的有界闭集.

再设, $a^0(\cdot):[t_0,\ t_1]\to\mathbb{R}^n$ 是可积的; $b^0(\cdot,\ \cdot):[t_0,\ t_1]\times\mathbb{R}^m\to\mathbb{R}$ 满足条件:

对于 $v\in\mathbb{R}^m$, $b^0(\cdot,\ v):[t_0,\ t_1]\to\mathbb{R}$ 是可积的;

对于 $t\in[t_0,\ t_1]$, $b^0(t,\ \cdot):\mathbb{R}^m\to\mathbb{R}$ 是连续的;

对于 $\mathbb{R}^m$ 中的任一有界集 K, 存在可积的 $\mu(\cdot):[t_0,\ t_1]\to\mathbb{R}$, 使得不等式

$$|b^0(t,\ v)|\leqslant\mu(t)$$

对于 $(t,\ v)\in[t_0,\ t_1]\times K$ 成立.

又设 x_0 和 x_1 是 $\mathbb{R}^n$ 中的两个给定点.

线性最优控制问题 寻求 $u(\cdot)\in U_{ad}$ 使得

$$x(t_0,\ u)=x_0,\quad x(t_1,\ u)=x_1,\tag{2}$$

并且 $$J(u(\cdot))=\int_{t_0}^{t_1}\{\langle a^0(t),\ x(t,\ u)\rangle+b^0(t,\ u(t))\}dt$$
$$\leqslant\int_{t_0}^{t_1}\{\langle a^0(t),\ x(t,\ v)\rangle+b^0(t,\ v(t))\}dt$$

对于 $v(\cdot)\in U_{ad}$, 且 $x(t_0,\ v)=x_0$, $x(t_1,\ v)=x_1$ 成立.

记 $$x^0(t)=\int_{t_0}^{t}\{\langle a^0(s),\ x(s,\ v)\rangle+b^0(s,\ v(s))\}ds,$$

那末 $$x^0(t_0,\ v)=0,\ x^0(t_1,\ v)=J(v(\cdot)),$$

$$\frac{dx^0(t,\ v)}{dt}=\langle a^0(t),\ x(t,\ v)\rangle+b^0(t,\ v(t)).$$

记 $\boldsymbol{x}=\begin{pmatrix} x^0 \\ x \end{pmatrix}$, $\boldsymbol{b}(t, v)=\begin{pmatrix} b^0(t, v) \\ b(t, v) \end{pmatrix}$, $\boldsymbol{x}_0=\begin{pmatrix} 0 \\ x_0 \end{pmatrix}$,

那末

$$\frac{d\boldsymbol{x}}{dt}=\begin{pmatrix} 0 & a^{0T}(t) \\ 0 & A(t) \end{pmatrix}\boldsymbol{x}+\begin{pmatrix} b^0(t, v(t)) \\ b(t, v(t)) \end{pmatrix}, \tag{3}$$

它关于 $\boldsymbol{x}$ 是线性的，边值条件成为

$$\boldsymbol{x}(t_0)=\boldsymbol{x}_0, \quad \boldsymbol{x}(t_1)=\begin{pmatrix} J(v(\cdot)) \\ x_1 \end{pmatrix}.$$

这样，最优控制问题可改写为：寻求 $u(\cdot)\in U_{ad}$，使得(3)的解 $\boldsymbol{x}(\cdot)=\boldsymbol{x}(\cdot, u)$，满足

$$\boldsymbol{x}(t_0)=\boldsymbol{x}_0,$$
$$x(t_1)=x_1,$$

并且 $x^0(t_1)\leqslant x^0(t_1, v)$ 对于 $v(\cdot)\in U_{ad}$，且

$$\boldsymbol{x}(t_0, v)=\boldsymbol{x}_0, \quad x(t_1, v)=x_1$$

成立。

若记系统(3)的等时区域为

$$\hat{R}(t)=\{\boldsymbol{x} \mid \boldsymbol{x}=\boldsymbol{x}(t, v),\ v(\cdot)\in U_{ad}\},$$

那末根据§6的定理3知，$\hat{R}(t)$ 是 $\mathbb{R}^{n+1}$ 中的有界凸闭集。

定理1 设 U 是 $\mathbb{R}^m$ 中的有界闭集，并且存在某一容许控制 $v(\cdot)$，使得 $x(t_0, v)=x_0$, $x(t_1, v)=x_1$，那末线性最优控制问题的解是存在的。

证 $\mathbb{R}^{n+1}$ 中的有界凸闭集 $\hat{R}(t_1)$ 与 $\mathbb{R}^{n+1}$ 中的直线

$$L: x=x_1$$

有公共点 $\boldsymbol{x}(t_1, v)$。集 $L\cap\hat{R}(t_1)$ 是 $\mathbb{R}^{n+1}$ 中的非空的有界凸闭集。

显然，线性最优控制问题等价于

$$\inf\{x^0 \mid \boldsymbol{x}\in L\cap\hat{R}(t_1)\}.$$

根据 $L\cap\hat{R}(t_1)$ 的闭性，得知存在 $u(\cdot)$，使得

$$x^0(t_1, u)=\inf\{x^0 | \boldsymbol{x}\in L\cap\hat{R}(t_1)\}.$$ ∎

定理 2 设 $u(\cdot)$ 是线性最优控制问题的解，$x(\cdot)=x(\cdot, u)$，那末存在 $\alpha_0\leqslant 0$ 和 $\psi(\cdot)$ 满足

$$\frac{d\psi(t)}{dt}=-A^T(t)\psi(t)-\alpha_0 a^0(t),$$

$$\alpha_0^2+|\psi(t_1)|^2>0,$$

使得最大值条件

$$\alpha_0 b^0(t, u(t))+\langle\psi(t), b(t, u(t))\rangle$$
$$=\max_{v\in U}\{\alpha_0 b^0(t, v)+\langle\psi(t), b(t, v)\rangle\}$$

在 $[t_0, t_1]$ 上几乎处处成立.

证 $(x^0(t_1, u), x_1)$ 是有界凸闭集 $\hat{R}(t_1)$ 的边界点，并且从它出发指向负 x^0 轴的射线 l:

$$l: x^0<x^0(t_1, u), \quad x=x_1$$

不在 $\hat{R}(t_1)$ 中.

根据凸集的分离性定理，在 $\mathbb{R}^{n+1}$ 中存在以 (α_0, λ) 为法向量的超平面，把 l 和 $\hat{R}(t_1)$ 分在两侧，即

$$\alpha_0^2+|\lambda|^2>0,$$

$$\alpha_0(-1)\geqslant 0, \tag{4}$$

$$\alpha_0(x^0(t_1, v)-x^0(t_1, u))+\langle\lambda, x(t_1, v)-x(t_1, u)\rangle\leqslant 0. \tag{5}$$

由(4)得到

$$\alpha_0\leqslant 0.$$

因为 $x(t, v)=\Phi(t, t_0)x_0+\int_{t_0}^t \Phi(t, s)b(s, v(s))ds,$

$$x(t, u)=\Phi(t, t_0)x_0+\int_{t_0}^t \Phi(t, s)b(s, u(s))ds,$$

所以得到

$$x(t, v)-x(t, u)=\int_{t_0}^{t}\Phi(t, s)\{b(s, v(s))-b(s, u(s))\}ds,$$

$$x^0(t_1, v)-x^0(t_1, u)=\int_{t_0}^{t_1}\{b^0(t, v(t))-b^0(t, u(t))\}dt$$

$$+\int_{t_0}^{t_1}\left\langle a^0(t), \int_{t_0}^{t}\Phi(t, s)\{b(s, v(s))-b(s, u(s))\}ds\right\rangle dt.$$

从而由(5)得到

$$\alpha_0\int_{t_0}^{t_1}b^0(t, v(t))dt+\int_{t_0}^{t_1}\int_{t_0}^{t}\langle\alpha_0a^0(t), \Phi(t, s)b(s, v(s))\rangle ds\, dt$$

$$+\int_{t_0}^{t_1}\langle\lambda, \Phi(t_1, t)b(t, v(t))\rangle dt$$

$$\leqslant\alpha_0\int_{t_0}^{t_1}b^0(t, u(t))dt$$

$$+\int_{t_0}^{t_1}\int_{t_0}^{t}\langle\alpha_0a^0(t), \Phi(t, s)b(s, u(s))\rangle ds\, dt$$

$$+\int_{t_0}^{t_1}\langle\lambda, \Phi(t_1, t)b(t, u(t))\rangle dt.$$

因此成立着

$$\alpha_0\int_{t_0}^{t_1}b^0(t, v(t))dt+\int_{t_0}^{t_1}\Big\langle\Phi^T(t_1, t)\lambda$$

$$+\alpha_0\int_{t}^{t_1}\Phi^T(\sigma, t)a^0(\sigma)d\sigma, b(t, v(t))\Big\rangle dt$$

$$\leqslant\alpha_0\int_{t_0}^{t_1}b^0(t, u(t))dt+\int_{t_0}^{t_1}\Big\langle\Phi^T(t_1, t)\lambda$$

$$+\alpha_0\int_{t}^{t_1}\Phi^T(\sigma, t)a^0(\sigma)d\sigma, b(t, u(t))\Big\rangle dt. \tag{6}$$

记 $$\psi(t)=\Phi^T(t_1, t)\lambda+\alpha_0\int_{t}^{t_1}\Phi^T(\sigma, t)a^0(\sigma)d\sigma,$$

那末 $\psi(\cdot)$ 适合

$$\frac{d\psi(t)}{dt}=-A^T(t)\psi(t)-\alpha_0a^0(t),$$

$$\psi(t_1)=\lambda,$$

所以 $$\alpha_0^2+\|\psi(t_1)\|^2=\alpha_0^2+\|\lambda\|^2>0.$$

不等式(6)成为：对于 $v(\cdot)\in U_{ad}$，成立着

$$\int_{t_0}^{t_1}\{\alpha_0 b^0(t,\ v(t))+\langle\psi(t),\ b(t,\ v(t))\rangle\}dt$$

$$\leqslant\int_{t_0}^{t_1}\{\alpha_0 b^0(t,\ u(t))+\langle\psi(t),\ b(t,\ u(t))\rangle\}dt \tag{7}$$

设 $\bar{t}$ 是 $\alpha_0 b^0(t,\ u(t))+\langle\psi(t),\ b(t,\ u(t))\rangle$ 的勒贝格点，即成立着

$$\lim_{h\to 0}\frac{1}{2h}\int_{\bar{t}-h}^{\bar{t}+h}\{\alpha_0 b^0(t,\ u(t))+\langle\psi(t),\ b(t,\ u(t))\rangle\}dt$$

$$=\alpha_0 b^0(\bar{t},\ u(\bar{t}))+\langle\psi(\bar{t}),\ b(\bar{t},\ u(\bar{t}))\rangle.$$

对于 $h>0$ 和 $v\in\mathbb{R}^m$，置

$$v(t)=\begin{cases} v, & \text{当 } |t-\bar{t}|\leqslant h \text{ 时},\\ u(t), & \text{当 } |t-\bar{t}|>h \text{ 时},\end{cases}$$

那末 $v(\cdot)\in U_{ad}$，并且由(7)得到

$$\int_{\bar{t}-h}^{\bar{t}+h}\{\alpha_0 b^0(t,\ v)+\langle\psi(t),\ b(t,\ v)\rangle\}dt$$

$$\leqslant\int_{\bar{t}-h}^{\bar{t}+h}\{\alpha_0 b^0(t,\ u(t))+\langle\psi(t),\ b(t,\ u(t))\rangle\}dt,$$

两端除以 $2h$，并令 $h\to 0$，得到

$$\alpha_0 b^0(\bar{t},\ v)+\langle\psi(\bar{t}),\ b(\bar{t},\ v)\rangle$$

$$\leqslant\alpha_0 b^0(\bar{t},\ u(\bar{t}))+\langle\psi(\bar{t}),\ b(\bar{t},\ u(\bar{t}))\rangle,$$

所以 $$\max_{v\in U}\{\alpha_0 b^0(\bar{t},\ v)+\langle\psi(\bar{t}),\ b(\bar{t},\ v)\rangle\}$$

$$=\alpha_0 b^0(\bar{t},\ u(\bar{t}))+\langle\psi(\bar{t}),\ b(\bar{t},\ u(\bar{t}))\rangle.$$

因为几乎所有的点是可积函数的勒贝格点，所以最大值条件几乎处处成立。■

习　　题

1　设 Ω 是 $\mathbb{R}^n$ 中的有界凸闭集，l 是 $\mathbb{R}^n$ 中的直线段，l 的点(除端点外)不是 Ω 的点，那末存在超平面把 Ω 和 l 分在两侧。

2　如果对于 $t\in[\alpha,\beta]$，$f(t,\cdot):\mathbb{R}^m\to\mathbb{R}$ 是连续的，对于 $u\in\mathbb{R}^m$，$f(\cdot,u):[\alpha,\beta]\to\mathbb{R}$ 是可积的，U 是 $\mathbb{R}^m$ 中的有界闭集，并且对于 $t\in[\alpha,\beta]$，有

$$\max_{v\in U} f(t,v)=\varphi(t),$$

$\varphi(\cdot)$ 在 $[\alpha,\beta]$ 上是可积的. 试证：存在 $u(\cdot):[\alpha,\beta]\to U$ 是可积的，使得

$$\varphi(\cdot)=f(\cdot,u(\cdot)).$$

3　利用上题及不等式(7)，证明最大值条件。

第三章　庞特里雅金的最大值原理

§1　最优控制问题的叙述和最大值原理

我们在第二章介绍了线性系统最优控制的一些结果，其中一个重要的结果是最大值原理．1956年，庞特里雅金和他的同事们正确地提出了最优控制问题的数学模型，并且以猜想的形式提出了解决最优控制问题的最大值原理；1958年他们首先公布了线性系统时间最优控制的最大值原理的证明；1960年他们刊出关于一般形式的最大值原理的证明．由于他们所提出的数学模型能确切地反映自动控制技术的发展，并且突破了古典变分学只能在开域中进行的框框，能够具体解决一般的时间最优控制问题，推广布绍的结果到一般情形，使庞特里雅金的最大值原理受到科学界的重视．

与此同时，贝尔曼用动态规划方法讨论最优控制问题，得到了人们称之为贝尔曼方程的必要条件．卡尔曼则具体研究了线性二次最优控制，建立了最优线性反馈调节器设计的理论基础(我们在第二章已讨论了)．1961年，伯科维茨用古典变分学方法推导出最大值原理．

这样，庞特里雅金的最大值原理、贝尔曼的动态规划方法、古典变分学的基本必要条件和卡尔曼的最优线性反馈调节器设计的理论，为解决最优控制问题奠定了理论基础．

现在我们来分析控制问题的正确叙述．

首先，设受控对象的状态是由下面的微分方程组

$$\frac{dx}{dt}=f(t,\ x,u(t)) \tag{1}$$

描述，其中 $f: \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的，并且 $\frac{\partial f}{\partial x}$ 也是连续的，这里

$$f=(f^1, f^2, \cdots, f^n)^T,$$

$$x=(x^1, x^2, \cdots, x^n)^T,$$

$$\frac{\partial f}{\partial x}=\begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^2}{\partial x^1} & \cdots & \frac{\partial f^n}{\partial x^1} \\ \frac{\partial f^1}{\partial x^2} & \frac{\partial f^2}{\partial x^2} & \cdots & \frac{\partial f^n}{\partial x^2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f^1}{\partial x^n} & \frac{\partial f^2}{\partial x^n} & \cdots & \frac{\partial f^n}{\partial x^n} \end{bmatrix}.$$

又设 U 是 $\mathbb{R}^m$ 中的给定集，它表征对控制器的限制，称为系统的控制区域。

如果 $u(\cdot): [t_0, t_1]\to\mathbb{R}^m$ 是有界可积的，并且在 $[t_0, t_1]$ 上几乎处处成立着 $u(t)\in U$，就称 $u(\cdot)$ 在 $[t_0, t_1]$ 上是系统(1)的容许控制．记为 $u(\cdot)\in U_{ad}$.

在第二章中，$f(t, x, u)=A(t)x+b(t, u)$，从而系统的状态方程是线性的．现在，状态方程是一般的非线性方程．

在控制问题中的微分方程组与常微分方程教程中所讨论的区别在于：这里有可供人们选取的控制作用 $u(\cdot)$，并且 $u(\cdot)$ 一般是受有限制或约束的，即它不能是任意的可积函数．同时，从绪论的例1看出，我们不能要求 $u(\cdot)$ 是连续函数．第二章中，我们已经看出，假设它为取值于 U 的有界可积函数或为逐段连续函数是适当的．

这样，对于确定的 $u(\cdot)\in U_{ad}$，方程组(1)的右端 $f(t, x, u(t))$ 关于 t 不可能是连续的，所以我们有必要讨论可积右端的微分方程组．既然这样，我们假设 f 满足下述条件：

对于 $x\in\mathbb{R}^n$ 和 $u\in\mathbb{R}^m$，$f(\cdot, x, u): [t_0, t_1]\to\mathbb{R}^n$ 是可积的，这

里 $t_0, t_1\in\mathbb{R}$;

对于 $t\in\mathbb{R}$, $f(t,\cdot,\cdot):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的;

对于 $(t,u)\in\mathbb{R}\times\mathbb{R}^m$, $\dfrac{\partial f(t,\cdot,u)}{\partial x}:\mathbb{R}^n\to\mathbb{R}^{n\times n}$ 是连续的.

在上述条件下, 对于 $u(\cdot)\in U_{ad}$, $f(\cdot,\cdot,u(\cdot))$: $[t_0,t_1]\times\mathbb{R}^n\to\mathbb{R}^n$ 满足第一章§3关于方程右端的条件. 从而, 如果给定 $(t_0,x_0)\in\mathbb{R}\times\mathbb{R}^n$, $u(\cdot)\in U_{ad}$, 那末方程组(1)以 (t_0,x_0) 为初值的解存在唯一. 自然, 由于方程组(1)是非线性的, 我们不能肯定它的解在整个 $[t_0,t_1]$ 上存在.

我们知道, 控制作用 $u(\cdot)$ 是为了使人们达到一定目的的.

设在 $\mathbb{R}^n$ 中还给定了点 x_1. 控制问题是选取容许控制 $u(\cdot)\in U_{ad}$, 使得系统(1)以 (t_0,x_0) 为初值的解 $x(\cdot)$ 在某一时刻 t_1 达到 x_1, 即

$$x(t_1)=x_1.$$

这时, 我们称容许控制 $u(\cdot)$ 把系统(1)从状态 x_0 迁移到状态 x_1. 这里时刻 t_1 没有预先给定, 以便包括第二章所讨论的时间最优控制问题. 当然, 也可以预先给定 t_1.

为了表征系统的性能, 我们引入性能指标

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,x(t),u(t))dt, \tag{2}$$

这里 f^0: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 适合类似于 f 的条件. 即对于 $(x,u)\in\mathbb{R}^n\times\mathbb{R}^m$, $f^0(\cdot,x,u)$: $[t_0,t_1]\to\mathbb{R}$ 是可积的; 对于 $t\in\mathbb{R}$, $f^0(t,\cdot,\cdot):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 是连续的; 对于 $(t,u)\in\mathbb{R}\times\mathbb{R}^m$, $\partial f^0(t,\cdot,u)/\partial x$: $\mathbb{R}^n\to\mathbb{R}^{n\times 1}$ 是连续的.

最优控制问题 寻求容许控制 $u(\cdot)\in U_{ad}$, 使得它把系统从状态 x_0 迁移到状态 x_1, 并且

$$\begin{aligned}J(u(\cdot))=\inf\{J(v(\cdot))\mid v(\cdot)\in U_{ad},\\ x(t_0,v)=x_0,\ x(t_1,v)=x_1\},\end{aligned} \tag{3}$$

这里 (t_0, x_0)，x_1 是给定的，而对于不同的 $v(\cdot)\in U_{ad}$，使得 $x(t_1, v)=x_1$ 成立的 t_1 可能是不同的.

如果 $u(\cdot)\in U_{ad}$ 满足(3)，就称 $u(\cdot)$ 为最优控制，$x(\cdot)=x(\cdot, u)$为最优轨线. 也称 $u(\cdot)$和 $x(\cdot)=x(\cdot, u)$为最优控制问题的解.

当然，该问题还可以推广，将在第四章详述.

为讨论该问题，可以引进一个新坐标 x^0：

$$\frac{dx^0}{dt}=f^0(t, x, u(t)),$$

$$x^0(t_0)=0.$$

若记

$$\boldsymbol{x}=\begin{pmatrix}x^0\\x\end{pmatrix},\quad \boldsymbol{f}=\begin{pmatrix}f^0\\f\end{pmatrix},\quad \boldsymbol{x}_0=\begin{pmatrix}0\\x_0\end{pmatrix},$$

那末可以考虑系统

$$\frac{d\boldsymbol{x}}{dt}=\boldsymbol{f}(t, x, u(t)),\ u(t)\in U, \boldsymbol{x}(t_0)=\boldsymbol{x}_0. \tag{4}$$

若记 L 为 $\mathbb{R}^{n+1}$ 中的直线

$$x=x_1,$$

那末最优控制问题可改写为：寻求 $u(\cdot)\in U_{ad}$，使得方程组(4)的解 $\boldsymbol{x}(\cdot, u)$在某一时刻 t_1 与 L 相交，即

$$\boldsymbol{x}(t_1, u)\in L,$$

并且 $x^0(t_1, u)$取最小值. 即，如果 $v(\cdot)\in U_{ad}$，使得 $\boldsymbol{x}(\cdot, v)$在某一时刻 $\tilde{t}_1$ 与 L 相交，即 $\boldsymbol{x}(\tilde{t}_1, v)\in L$，那末成立着

$$x^0(t_1, u)\leqslant x^0(\tilde{t}_1, v).$$

引进哈密顿函数

$$\mathscr{H}(t, \psi, x, u)=\psi_0 f^0(t, x, u)+\langle\psi, f(t, x, u)\rangle,$$

那末系统(4)可表示为

$$\frac{d\boldsymbol{x}}{dt}=\frac{\partial\mathscr{H}(t, \boldsymbol{\psi}, x, u(t))}{\partial\boldsymbol{\psi}},\quad \boldsymbol{x}(t_0)=\boldsymbol{x}_0.$$

这里 $$\boldsymbol{\psi}=\begin{pmatrix}\psi_0\\ \psi\end{pmatrix},\quad \psi=(\psi_1,\ \cdots,\ \psi_n)^T.$$

庞特里雅金的最大值原理 设 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,\ u)$ 是最优控制问题的解，那末存在 $\psi_0\leqslant 0$ 和 $\psi(\cdot)$，满足

$$\psi_2^0+\|\psi(t)\|^2>0, \tag{5}$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial f(t,x(t),u(t))}{\partial x}\psi(t)-\psi_0\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}, \tag{6}$$

使得最大值条件

$$\mathscr{H}(t,\ \boldsymbol{\psi}(t),\ x(t),\ u(t))=\max_{v\in U}\mathscr{H}(t,\ \boldsymbol{\psi}(t),\ x(t),\ v) \tag{7}$$

在 $[t_0,\ t_1]$ 上几乎处处成立。

如果 t_1 不是预先给定的，那末

$$\begin{aligned}\mathscr{H}(t_1,\ \boldsymbol{\psi}(t_1),\ x(t_1),\ u(t_1))&=\max_{v\in U}\mathscr{H}(t_1,\ \boldsymbol{\psi}(t_1),\ x(t_1),\ v)\\&=0.\end{aligned} \tag{8}$$

第四章将给出最大值原理的严格证明。

方程组(6)称为最优控制问题的伴随方程组，它可以写为

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(t,\ \boldsymbol{\psi}(t),\ x(\cdot),\ u(t))}{\partial x}.$$

当 $f^0\equiv 1$，t_1 不预先给定时，性能指标

$$J(u(\cdot))=t_1-t_0$$

表示控制作用 $u(\cdot)$ 把系统从状态 x_0 迁移到 x_1 的时间。这时，哈密顿函数 $\mathscr{H}$ 可表示为

$$\begin{aligned}\mathscr{H}(t,\ \boldsymbol{\psi},\ x,\ u)&=\psi_0+\langle\psi,\ f(t,\ x,\ u)\rangle\\&=\psi_0+H(t,\ \psi,\ x,\ u),\end{aligned}$$

伴随方程组成为

$$\frac{d\psi}{dt}=-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\psi, \tag{9}$$

并且 $$\|\psi(t)\|^2>0.$$
否则，由(8)推出 $\psi_0=0$，它与 $\psi_0^2+\|\psi(t)\|^2>0$ 矛盾。

因此，对于时间最优控制问题的最大值原理是：设 $u(\cdot)$ 和 $x(\cdot)=x(\cdot, u)$ 是时间最优控制问题的解，那末存在方程组(9)的非平凡解 $\psi(\cdot)$，使得最大值条件

$$H(t, \psi(t), x(t), u(t))=\max_{v\in U} H(t, \psi(t), x(t), v) \tag{10}$$

在 $[t_0, t_1]$ 上几乎处处成立。

特别，对于线性系统

$$\frac{dx}{dt}=A(t)x+b(t, u(t)), \quad u(t)\in U,$$
$$x(t_0)=x_0,$$

最快把状态 x_0 迁移到 x_1 的时间最优控制问题，哈密顿函数是

$$H(t, \psi, x, u)=\langle\psi, A(t)x+b(t, u)\rangle,$$

伴随方程组是

$$\frac{d\psi}{dt}=-A^T(t)\psi, \tag{11}$$

它与控制作用 $u(\cdot)$ 及轨线 $x(\cdot)$ 无关，是可以直接解出的。最大值原理是说：存在(11)的非平凡解 $\psi(\cdot)$，使得最大值条件

$$\langle\psi(t), A(t)x(t)+b(t, u(t))\rangle$$
$$=\max_{v\in U}\langle\psi(t), A(t)x(t)+b(t, v)\rangle$$

在 $[t_0, t_1]$ 上几乎处处成立。它就是

$$\langle\psi(t), b(t, u(t))\rangle=\max_{v\in U}\langle\psi(t), b(t, v)\rangle$$

在 $[t_0, t_1]$ 上几乎处处成立。这正是第二章§6的结论。在第二章中，还给出了最优时间的计算公式和 $\psi(t_1)$ 如何确定的公式。

习　题

1　试用最大值原理来讨论第二章§8的最优控制问题。

2 试用最大值原理来讨论线性二次最优控制问题的必要条件。

§2 动态规划方法与最大值原理

贝尔曼的动态规划方法能用来研究很广泛的最优化问题，这里仅就§1所讨论的最优控制问题来讨论动态规划方法.

考虑系统

$$\frac{dx}{dt}=f(t,\ x,\ u(t)),\quad u(t)\in U,\tag{1}$$

其中 f: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的, $\frac{\partial f}{\partial x}$: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{n\times n}$ 是连续的，容许控制 $u(\cdot)$ 是取值于 U 的具有有限个第一类间断点的逐段连续函数，记为 $u(\cdot)\in U_{ad}$. 性能指标为泛函

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,\ x(t),\ u(t))dt,\tag{2}$$

这里 f^0 有与 f 相同的连续性.

最优控制问题 设给定 $x_1\in\mathbb{R}^n$. 对于 $(\tau,\ \xi)\in\mathbb{R}\times\mathbb{R}^n$，如果存在 $v(\cdot)\in U_{ad}$，使得(1)以 $(\tau,\ \xi)$ 为初值的解 $x=x(\cdot,\ v)$ 满足

$$x(\tau,\ v)=\xi,\quad x(t_1,\ v)=x_1,$$

就称 $v(\cdot)$ 把系统(1)从初值 $(\tau,\ \xi)$ 迁移到状态 x_1. 最优控制问题是：在把系统(1)从初值 $(\tau,\ \xi)$ 迁移到状态 x_1 的容许控制中，选取使泛函(2)达到最小的 $u(\cdot)\in U_{ad}$，即

$$x(\tau,\ u)=\xi,\quad x(t_1,\ u)=x_1,$$

而

$$\begin{aligned}J(u(\cdot))=\min\{J(v(\cdot))\,|\,&x(\tau,\ v)=\xi,\\ &x(t_1,\ v)=x_1,\ v(\cdot)\in U_{ad}\}.\end{aligned}\tag{3}$$

显然，对于 $(\tau,\ \xi)\in\mathbb{R}\times\mathbb{R}^n$，(3)有依赖于 $(\tau,\ \xi)$ 的值，记它为 $S(\tau,\ \xi)$，称为最优控制问题(3)的值函数. 有时，为强调(3)依赖于初值 $(\tau,\ \xi)$，称它为以 $(\tau,\ \xi)$ 为初值的最优控制问题. 如果对于

(τ, ξ)，没有容许控制把系统(1)从初值(τ, ξ)迁移到x_1，就认为

$$S(\tau, \xi)=+\infty.$$

贝尔曼运用动态规划方法研究了最优控制问题，它的理论基础是下面的原理.

最优性原理 如果一个系统从时刻t_A到时刻t_B的状态变化具有某种最优性质，t_C是位于t_A和t_B之间的时刻，那末系统从时刻t_C到时刻t_B的状态变化也具有这种最优性质.

对于一般情形的最优性原理是毋需证明的. 对于最优控制问题，最优性原理是下面的：如果$u(\cdot)$和$x(\cdot)=x(\cdot, u)$是以(τ, ξ)为初值的最优控制问题(3)的解，$t'\in(\tau, t_1)$，$x'=x(t', u)$，那末$u(\cdot)$和$x(\cdot)$也是以(t', x')为初值的最优控制问题的解.

事实上，如果存在$v(\cdot)\in U_{ad}$，使得(1)以(t', x')为初值的解$x(\cdot, v)$满足

$$x(t', v)=x', \quad x(t_1', v)=x_1,$$

并且

$$\int_{t'}^{t_1'} f^0(t, x(t, v), v(t))dt<\int_{t'}^{t_1} f^0(t, x(t), u(t))dt.$$

规定$v(\cdot)$在$[\tau, t']$就是$u(\cdot)$，那末$v(\cdot)$在$[\tau, t_1']$上是容许控制，并且

$$x(\tau, v)=x(\tau, u)=\xi, \quad x(t', v)=x(t', u)=x',$$
$$x(t_1', v)=x_1,$$

$$\int_{\tau}^{t_1'} f^0(t, x(t, v), v(t))dt$$
$$=\int_{\tau}^{t'} f^0(t, x(t), u(t))dt+\int_{t'}^{t_1'} f^0(t, x(t, v), v(t))dt,$$

上式右端的第二项小于

$$\int_{t'}^{t_1} f^0(t, x(t), u(t))dt,$$

所以

$$\int_{\tau}^{t'} f^0(t,\ x(t,\ v),v(t))dt<\int_{\tau}^{t_1} f^0(t,\ x(t),\ u(t))dt.$$

这与$u(\cdot)$的最优性矛盾. 因此, $u(\cdot)$ 和 $x(\cdot)$是以$(t',\ x')$ 为初值的最优控制问题的解.

如果称$[t',\ t_1]$上的轨线 $x(\cdot)$为第二段,那末最优性原理可表述为:最优轨线的第二段仍是最优轨线.

一　贝尔曼的动态规划方程

现推导最优控制问题的值函数$S(\tau,\ \xi)$所满足的关系式. 假设在$\mathbb{R}\times\mathbb{R}^n$ 的某一区域中, $S(\cdot,\ \cdot)$是连续可微的. $u(\cdot)$和$x(\cdot)=x(\cdot,\ u)$是以$(\tau,\ \xi)$为初值的最优控制问题的解,即

$$\begin{aligned}S(\tau,\ \xi)&=\int_{\tau}^{t_1} f^0(t,\ x(t),\ u(t))dt\\&=\min\{J(v(\cdot))\,|\,x(\tau,\ v)=\xi,\\&\qquad x(t_1',\ v)=x_1,\ v(\cdot)\in U_{ad}\}.\end{aligned}$$

根据最优性原理,$u(\cdot)$和 $x(\cdot)$也是以$(t,\ x(t))$为初值的最优控制问题的解,所以

$$S(t,\ x(t))\equiv\int_{t}^{t_1} f^0(s,\ x(s),\ u(s))ds. \tag{4}$$

两端对 t 求导,得到

$$\frac{d}{dt}S(t,\ x(t))\equiv-f^0(t,\ x(t),\ u(t)),$$

即

$$\frac{\partial S(t,\ x(t))}{\partial t}+\left\langle\frac{\partial S(t,\ x(t))}{\partial x},\ \frac{dx(t)}{dt}\right\rangle\equiv-f^0(t,\ x(t),\ u(t)),$$

或

$$\begin{aligned}f^0(t,\ x(t),\ u(t))+\left\langle\frac{\partial S(t,\ x(t))}{\partial x},\ f(t,\ x(t),\ u(t))\right\rangle&\\+\frac{\partial S(t,\ x(t))}{\partial t}\equiv0,&\end{aligned} \tag{5}$$

这里
$$\frac{\partial S}{\partial x}=\left(\frac{\partial S}{\partial x^1},\ \frac{\partial S}{\partial x^2},\ \cdots,\ \frac{\partial S}{\partial x^n}\right)^T.$$

特别，对于 $t=\tau$, $x(\tau)=\xi$, (5) 成为

$$f^0(\tau,\ \xi,\ u(\tau))+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial \xi},\ f(\tau,\ \xi,\ u(\tau))\right\rangle+\frac{\partial S(\tau,\ \xi)}{\partial \tau}=0. \tag{6}$$

又设在 $[\tau,\ \tau+\Delta\tau]$ 上，$v(t)=v$, $v\in\mathbb{R}^m$，方程 (1) 以 $(\tau,\ \xi)$ 为初值的解 $x(\cdot,\ v)$，即

$$x(\tau,\ v)=\xi,$$

它在 $\tau+\Delta\tau$ 处为 $x(\tau+\Delta\tau,\ v)$.

设以 $(\tau+\Delta\tau,\ x(\tau+\Delta\tau,\ v))$ 为初值的最优控制问题的解为 $\tilde{u}(\cdot)$ 和 $\tilde{x}(\cdot)=x(\cdot,\ \tilde{u})$，即存在 t_1'，使得

$$x(t_1',\ \tilde{u})=x_1,$$

$$x(\tau+\Delta\tau,\ \tilde{u})=x(\tau+\Delta\tau,\ v),$$

$$S(\tau+\Delta\tau,\ \tilde{x}(\tau+\Delta\tau))=\int_{\tau+\Delta\tau}^{t_1'} f^0(t,\ \tilde{x}(t),\ \tilde{u}(t))dt.$$

对于 $t\in[\tau,\tau+\Delta\tau)$，置 $\tilde{u}(t)=v$，那末 $\tilde{u}(\cdot)$ 在 $[\tau,\ t_1']$ 上是容许控制，并且

$$x(\tau,\ \tilde{u})=\xi,\qquad x(\tau+\Delta\tau,\ \tilde{u})=x(\tau+\Delta\tau,\ v),$$
$$x(t_1',\ \tilde{u})=x_1,$$

所以

$$\begin{aligned}S(\tau,\ \xi)&\leqslant\int_{\tau}^{\tau+\Delta\tau} f^0(t,\ x(t,\ v),\ v)dt\\&\quad+\int_{\tau+\Delta\tau}^{t_1'} f^0(t,\ \tilde{x}(t),\ \tilde{u}(t))dt.\\&=\int_{\tau}^{\tau+\Delta\tau} f^0(t,\ x(t,\ v),\ v)dt\\&\quad+S(\tau+\Delta\tau,\ x(\tau+\Delta\tau,\ v)).\end{aligned}$$

因此

$$0\leqslant\frac{1}{\Delta\tau}\int_{\tau}^{\tau+\Delta\tau}f^0(t,\ x(t,\ v),\ v)dt$$
$$+\frac{S(\tau+\Delta\tau,\ x(\tau+\Delta\tau,\ v))-S(\tau,\ \xi)}{\Delta\tau}.$$

令 $\Delta\tau\to 0$, 由上式得到

$$0\leqslant f^0(\tau,\ \xi,\ v)+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial\xi},\ f(\tau,\ \xi,\ v)\right\rangle+\frac{\partial S(\tau,\ \xi)}{\partial\tau}.$$

它与(6)相结合，就得到

$$0=f^0(\tau,\ \xi,\ u(\tau))+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial\xi},\ f(\tau,\ \xi,\ u(\tau))\right\rangle$$
$$+\frac{\partial S(\tau,\ \xi)}{\partial\tau}$$
$$=\min_{v\in U}\left\{f^0(\tau,\ \xi,\ v)+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial\xi},\ f(\tau,\ \xi,\ v)\right\rangle\right.$$
$$\left.+\frac{\partial S(\tau,\ \xi)}{\partial\tau}\right\}.$$

因为 $\partial S(\tau,\ \xi)/\partial\tau$ 与 v 无关，所以由上式得到

$$-\frac{\partial S(\tau,\ \xi)}{\partial\tau}$$
$$=f^0(\tau,\ \xi,\ u(\tau))+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial\xi},\ f(\tau,\ \xi,\ u(\tau))\right\rangle$$
$$=\min_{v\in U}\left\{f^0(\tau,\ \xi,\ v)+\left\langle\frac{\partial S(\tau,\ \xi)}{\partial\xi},\ f(\tau,\ \xi,\ v)\right\rangle\right\}. \tag{7}$$

方程(7)称为贝尔曼的动态规划方程，由于在(7)中出现求最小值，所以它是一个非线性偏微分方程.

显然，对于以 $(\tau,\ x_1)$ 为初值的控制问题，$t_1=\tau$，最优控制的值函数应取值 0，即

$$S(\tau,\ x_1)=0, \tag{8}$$

这是(7)的边值条件.

应当指出的是，在上述推导中，我们假设了 $S(\cdot,\ \cdot)$ 的一阶连

续可微性.

二 从动态规划方程推导最大值原理

关于最大值原理的严格证明,将在第四章进行.现在在$S(\cdot,\cdot)$两阶连续可微的条件下,由动态规划方程来推导最大值原理.

设$u(\cdot)$和$x(\cdot)=x(\cdot, u)$是以(t_0, x_0)为初值的最优控制问题的解,那末它也是以$(t, x(t))$为初值的最优控制问题的解,由(7)得到

$$f^0(t,x(t),u(t))+\left\langle\frac{\partial S(t, x(t))}{\partial x}, f(t, x(t), u(t))\right\rangle$$
$$=\min_{v\in U}\left\{f^0(t, x(t), v)+\left\langle\frac{\partial S(t, x(t))}{\partial x}, f(t, x(t), v)\right\rangle\right\}.$$

置

$$\psi_0=-1, \quad \psi_i(t)\equiv-\frac{\partial S(t, x(t))}{\partial x^i}, \quad i=1, 2, \cdots, n, \tag{9}$$

那末由上式得到

$$\psi_0 f^0(t, x(t), u(t))+\langle\psi(t), f(t, x(t), u(t))\rangle$$
$$=\max_{v\in U}\{\psi_0 f^0(t, x(t), v)+\langle\psi(t), f(t, x(t), v)\rangle\}. \tag{10}$$

这就是§1中最大值条件.

剩下应证明:$\psi(\cdot)$适合§1的伴随方程.

由(9)推出

$$\frac{d\psi_i(t)}{dt}=-\left\{\frac{\partial^2 S(t, x(t))}{\partial t\,\partial x^i}\right.$$
$$\left.+\sum_{j=1}^{n}\frac{\partial^2 S(t, x(t))}{\partial x^j\,\partial x^i}f^j(t, x(t), u(t))\right\},$$
$$(i=1, 2, \cdots, n). \tag{11}$$

对$(t, x)\in\mathbb{R}\times\mathbb{R}^n$,记

$$g(t, x)=f^0(t, x, u(t))+\left\langle\frac{\partial S(t, x)}{\partial x}, f(t, x, u(t))\right\rangle$$

$$+\frac{\partial S(t, x)}{\partial t}.$$

它当 $x=x(t)$ 时达到最小值 0. 因此

$$0=\frac{\partial g(t, x(t))}{\partial x^i}$$

$$=\frac{\partial f^0(t, x(t), u(t))}{\partial x^i}+\sum_{j=1}^{n}\frac{\partial S(t, x(t))}{\partial x^j}\frac{\partial f^j(t, x(t), u(t))}{\partial x^i}$$

$$+\sum_{j=1}^{n}\frac{\partial^2 S(t, x(t))}{\partial x^j\partial x^i}f^j(t, x(t), u(t))+\frac{\partial^2 S(t, x(t))}{\partial t\,\partial x^i},$$

把上式代入(11),得到

$$\frac{d\psi_i(t)}{dt}=\frac{\partial f^0(t, x(t), u(t))}{\partial x^i}-\sum_{j=1}^{n}\frac{\partial f^j(t, x(t), u(t))}{\partial x^i}\psi_j(t),$$

$$(i=1, 2, \cdots, n),$$

即

$$\frac{d\psi(t)}{dt}=-\psi_0\frac{\partial f^0(t, x(t), u(t))}{\partial x}-\frac{\partial f(t, x(t), u(t))}{\partial x}\psi(t).$$

它就是 §1 的伴随方程.

因此,当 $S(\cdot, \cdot)$ 二阶连续可微时,最大值原理是成立的.

又因为(8)成立,所以

$$\frac{\partial S(\tau, x_1)}{\partial \tau}=0.$$

根据(7)和(9),我们从上式推出

$$\psi_0 f^0(t_1, x_1, u(t_1))+\langle\psi(t_1), f(t_1, x_1, u(t_1))\rangle$$

$$=\frac{\partial S(t_1, x_1)}{\partial \tau}=0.$$

这就是 §1 中关于确定 t_1 的方程成立.

记

$$\mathscr{H}(t, \psi, x, v)=\psi_0 f^0(t, x, v)+\langle\psi, f(t, x, v)\rangle,$$

那末成立着

$$\mathscr{H}(t, \psi(t), x(t), u(t))=\max_{v\in U}\mathscr{H}(t, \psi(t), x(t), v),$$

$$\mathscr{H}(t_1,\ \psi(t_1),\ x(t_1),\ u(t_1))=0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))}{\partial x}.$$

近年来，人们对动态规划方程解的概念进行了研究，提出了方程(7)的粘性解的理论。

习　　题

1　对于系统

$$\frac{dx}{dt}=y,\quad \frac{dy}{dt}=u(t),\ |u(t)|\leqslant 1,$$

试求出把点$(\xi,\ \eta)$迁移到$(0,\ 0)$的时间最优控制问题的值函数$S(\tau,\ \xi,\ \eta)$。并指出$S(\tau,\ \xi,\ \eta)=S(0,\ \xi,\ \eta)$，并且它在$\eta^2=2\xi(\eta\leqslant 0)$上不是连续可微的。

2　试推导时间最优控制问题的值函数$T(\xi)$所适合的动态规划方程。

3　设$Q\subset\mathbb{R}^n$是给定的目标集，试用动态规划方法讨论$x(t_1,\ v)\in Q$的最优控制问题。导出动态规划方程和最大值原理(假定值函数是充分光滑的)，并导出最优轨线$x(\cdot)$的终点$x(t_1)\in Q$所满足的条件。

§3　最大值原理与古典变分学

我们知道，最优控制问题是讨论泛函极值的，它本质上是变分学所讨论的问题。

我们要指出，从最大值原理能推出古典变分学中的一切必要条件。并指出，最优控制问题是古典变分学中的拉格朗日(J．L．Lagrange)问题的推广；当控制区域U是$\mathbb{R}^m$中开集时，它们是等价的。但是，在最优控制问题中，U是$\mathbb{R}^m$中的任一集合(特别包括U是有界闭集的情形)，这时最大值原理是适用的，但古典变分学的方法发生了困难。

一　变分学的基本问题

设$x=x(\cdot):[t_0,\ t_1]\to\mathbb{R}^n$是连续且逐段有连续导数的，就称

曲线 $x=x(\cdot)$ 是逐段光滑的.

对于给定的 $(t_0, x_0)\in\mathbb{R}\times\mathbb{R}^n$ 和 $(t_1, x_1)\in\mathbb{R}\times\mathbb{R}^n$, 如果

$$x(t_0)=x_0,\quad x(t_1)=x_1,$$

我们称曲线连结点(t_0, x_0)和(t_1, x_1).

对于适合

$$\|\tilde{x}(t)-x(t)\|<\delta,\quad t\in[t_0, t_1].$$

的逐段光滑且 $\tilde{x}(t_0)=x_0$, $\tilde{x}(t_1)=x_1$ 的 $\tilde{x}(\cdot)$ 的全体, 称为曲线 $x(\cdot)$ 的 δ-邻域.

设 $G\subset\mathbb{R}^n$ 是开集, f: $[t_0, t_1]\times G\times\mathbb{R}^n\to\mathbb{R}$ 是连续可微的, x_0, $x_1\in G$. 对于逐段光滑的 $x(\cdot)$: $[t_0, t_1]\to G$, 考虑泛函

$$J(x)=\int_{t_0}^{t_1}f\left(t,\ x(t),\ \frac{dx(t)}{dt}\right)dt. \tag{1}$$

变分学问题 在连结点 (t_0, x_0) 和 (t_1, x_1) 的逐段光滑曲线 $x(\cdot)$的 δ-邻域中, 寻找使泛函(1)取极小值的曲线.

如果 $x=x(\cdot)$ 使泛函(1)达到极小值, 就称它是变分学问题的极值曲线.

变分学的基本问题是求出它的全部极值曲线.

考虑微分方程组

$$\frac{dx}{dt}=u(t), \tag{2}$$

这里 $u(\cdot)$: $[t_0, t_1]\to\mathbb{R}^n$ 是逐段连续的, 因而是有界可积的.

初值条件为

$$x(t_0)=x_0, \tag{3}$$

终值条件为

$$x(t_1)=x_1. \tag{4}$$

性能指标为使下述泛函取最小值:

$$J(x, u)=\int_{t_0}^{t_1}f(t,\ x(t),\ u(t))dt. \tag{5}$$

显然，如果 $x(\cdot)$ 是变分学问题的极值曲线，那末 $u(\cdot)$ 和 $x(\cdot)$ 必是最优控制问题(2)，(3)，(4)，(5)的解．反之，如果 $u(\cdot)$ 和 $x(\cdot)$ 是上述最优控制问题的解，那末 $x(\cdot)$ 是变分学问题(1)的解．

现在用最大值原理来推导变分学问题(1)的极值曲线所满足的必要条件．置

$$\mathscr{H}=\psi_0 f(t, x, v)+\langle\psi, v\rangle,$$

根据最大值原理，存在 $\psi_0\leqslant 0$ 和 $\psi(\cdot)$ 适合下述的伴随方程组

$$\frac{d\psi}{dt}=-\psi_0\frac{\partial f(t, x(t), u(t))}{\partial x}. \tag{6}$$

因为，在变分学问题的等价最优控制问题中，控制区域 U 是 $\mathbb{R}^n$，所以由最大值条件推出

$$\left.\frac{\partial\mathscr{H}(t, \psi(t), x(t), u)}{\partial u}\right|_{u=u(t)}=0,$$

即

$$\psi_0\frac{\partial f(t, x(t), u(t))}{\partial u}+\psi(t)\equiv 0. \tag{7}$$

由此推知 $\psi_0\neq 0$．否则，如果 $\psi_0=0$，由上式得出 $\psi(t)\equiv 0$，它与 $\psi_0^2+\|\psi(t)\|^2>0$ 相矛盾．

取 $\psi_0=-1$，(6)成为

$$\frac{d\psi(t)}{dt}=\frac{\partial f(t, x(t), u(t))}{\partial x}, \tag{8}$$

而(7)成为

$$\psi(t)\equiv\frac{\partial f(t, x(t), u(t))}{\partial u}. \tag{9}$$

由(8)积分得到

$$\psi(t)\equiv\psi(t_0)+\int_{t_0}^{t}\frac{\partial f(s, x(s), u(s))}{\partial x}ds,$$

由上式和(9)得到

$$\frac{\partial f(t,\ x(t),\ u(t))}{\partial u} \equiv c + \int_{t_0}^{t} \frac{\partial f(s,\ x(s),\ u(s))}{\partial x}\, ds,$$

这里 $c=\psi(t_0)$ 是常向量. 如果注意到(2), 我们由上式推出

$$\frac{\partial f\left(t,\ x(t),\ \dfrac{dx(t)}{dt}\right)}{\partial u} \equiv c + \int_{t_0}^{t} \frac{\partial f\left(s,\ x(s),\ \dfrac{dx(s)}{ds}\right)}{\partial x}\, ds. \tag{10}$$

这是积分形式的欧拉(L. Euler)方程.

当 f 和 x 二阶连续可微时, 由上式得到通常形式的欧拉方程

$$\frac{\partial f\left(t,\ x(t),\ \dfrac{dx(t)}{dt}\right)}{\partial x} - \frac{d}{dt}\left(\frac{\partial f\left(t,\ x(t),\ \dfrac{dx(t)}{dt}\right)}{\partial u}\right) = 0.$$

设 $f(t,\ x,\ \cdot): \mathbb{R}^n \to \mathbb{R}^n$ 是二阶连续可微的. 因为

$$\mathscr{H}(t,\ \psi(t),\ x(t),\ v)$$

在 $v=u(t)$ 达到最大值, 所以二次型

$$\sum_{i,j=1}^{n} \frac{\partial^2 \mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))}{\partial u^i\, \partial u^j} \xi^i \xi^j \leqslant 0,$$

从而得到

$$\sum_{i,j=1}^{n} \frac{\partial^2 f(t,\ x(t),\ \dot{x})}{\partial u^i\, \partial u^j} \xi^i \xi^j \geqslant 0. \tag{11}$$

这是曲线 $x(\cdot)$ 成为积分(1)的极值曲线的一个必要条件, 称为勒让特(A. M. Legendre)必要条件.

这样, 我们得到

定理 1 设 $x(\cdot)$ 是变分学问题(1)的极值曲线, 则沿着极值曲线 $x=x(\cdot)$ 适合 Euler 方程

$$\frac{d}{dt}\left(\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial \dot{x}}\right) = \frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial x}$$

及适合 Legendre 必要条件

$$\sum_{i,j=1}^{n} \frac{\partial^2 f(t, x(t), \dot{x}(t))}{\partial \dot{x}^i\, \partial \dot{x}^j} \xi^i \xi^j \geqslant 0.$$

二　变分学的拉格朗日问题

设 $G\subset\mathbb{R}^n\times\mathbb{R}^m$ 是开集. $f:[t_0,t_1]\times G\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续可微的; $h:[t_0,t_1]\times G\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 是连续可微的。

$(x_0, y_0)\in G$ 和 $(x_1, y_1)\in G$ 给定.

逐段光滑曲线

$$x=x(\cdot),\quad y=y(\cdot),$$

如果对于 $t\in[t_0, t_1]$, 成立着

$$(x(t), y(t))\in G,$$

并且

$$\begin{aligned} x(t_0)&=x_0, & y(t_0)&=y_0,\\ x(t_1)&=x_1, & y(t_1)&=y_1, \end{aligned}\tag{12}$$

就称它连结点 (x_0, y_0) 和 (x_1, y_1). 如果它还满足

$$\frac{dx(t)}{dt}-f\left(t, x(t), y(t), \frac{dy(t)}{dt}\right)\equiv 0,\tag{13}$$

就称它是容许曲线.

拉格朗日问题　求满足条件(12)和微分方程组(13)的容许曲线, 使得泛函

$$J(x, y)=\int_{t_0}^{t_1}h\left(t, x(t), y(t), \frac{dx(t)}{dt}, \frac{dy(t)}{dt}\right)dt\tag{14}$$

取最小值.

使(14)取最小值的容许曲线, 称为拉格朗日问题的极值曲线。

如果引入 $\dot{y}=u$, 考虑控制过程

$$\frac{dx}{dt}=f(t, x, y, u(t)),\qquad \frac{dy}{dt}=u(t),\tag{15}$$

边值条件仍取为(12). 置

$$f^0(t, x, y, u)=h(t, x, y, f(t, x, y, u), u),$$

考虑性能指标

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t, x(t), y(t), u(t))dt.\tag{16}$$

最优控制问题　选取逐段连续的 $u(\cdot):[t_0, t_1]\to\mathbb{R}^m$，使得方程组(15)的解 $x=x(\cdot)$，$y=y(\cdot)$ 满足边值条件(12)，并且使得泛函(16)取最小值.

显然，如果 $x(\cdot)$，$y(\cdot)$ 是拉格朗日问题的极值曲线，$u(t)=\dot{y}(t)$，那末 $u(\cdot)$：$[t_0, t_1]\to\mathbb{R}^m$ 是逐段连续的，并且 $u(\cdot)$、$x(\cdot)$ 和 $y(\cdot)$ 适合(15)和边值条件(12)，它使泛函(16)取最小值. 反之，如果 $u(\cdot)$，$x(\cdot)$ 和 $y(\cdot)$ 适合(15)和(12)，那末 $x(\cdot)$ 和 $y(\cdot)$ 适合(13)和(12). 并且如果 $u(\cdot)$，$x(\cdot)$ 和 $y(\cdot)$ 是最优控制问题的解，那末 $x(\cdot)$ 和 $y(\cdot)$ 是拉格朗日问题的极值曲线.

现在用最大值原理来推导拉格朗日问题的极值曲线所满足的必要条件.

设 $u(\cdot)$，$x(\cdot)$ 和 $y(\cdot)$ 是最优控制问题的解. 引进哈密顿函数

$$\mathscr{H}(t, \psi, x, y, u)=\psi_0 f^0(t, x, y, u)+\sum_{i=1}^{n}\psi_i f^i(t, x, y, u)+\sum_{j=1}^{m}\psi_{n+j}u^j. \tag{17}$$

这时，相应的伴随方程组是

$$\begin{aligned}\frac{d\psi_i}{dt}&=-\frac{\partial\mathscr{H}}{\partial x^i}\\&=-\psi_0\frac{\partial f^0(t, x(t), y(t), u(t))}{\partial x^i}\\&\quad-\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t, x(t), y(t), u(t))}{\partial x^i}\psi_\alpha(t),\qquad i=1,\cdots,n,\\\frac{d\psi_{n+j}}{dt}&=-\frac{\partial\mathscr{H}}{\partial y^j}\\&=-\psi_0\frac{\partial f^0(t, x(t), y(t), u(t))}{\partial y^j}\\&\quad-\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t, x(t), y(t), u(t))}{\partial y^j}\psi_\alpha(t),\qquad j=1,\cdots,m.\end{aligned} \tag{18}$$

再注意到$U=\mathbb{R}^m$，由最大值条件推知，在$[t_0, t_1]$上几乎处处成立着

$$0=\frac{\partial\mathscr{H}(t, \psi(t), x(t), y(t), u(t))}{\partial u^j}$$

$$=\psi_0\frac{\partial f^0(t, x(t), y(t), u(t))}{\partial u^j}$$

$$+\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t, x(t), y(t), u(t))}{\partial u^j}\psi_\alpha(t)+\psi_{n+j}(t),$$

$$(j=1, 2, \cdots, m). \tag{19}$$

根据

$$f^0(t, x, y, u)=h(t, x, y, f(t, x, y, u), u),$$

所以得到

$$\left.\begin{aligned}\frac{\partial f^0}{\partial u^j}&=\frac{\partial h}{\partial \dot{x}}\frac{\partial f}{\partial u^j}+\frac{\partial h}{\partial u^j},\\ \frac{\partial f^0}{\partial y^j}&=\frac{\partial h}{\partial \dot{x}}\frac{\partial f}{\partial y^j}+\frac{\partial h}{\partial y^j},\end{aligned}\right\}\quad (j=1, 2, \cdots, m). \tag{20}$$

$$\frac{\partial f^0}{\partial x^i}=\frac{\partial h}{\partial \dot{x}}\frac{\partial f}{\partial x^i}+\frac{\partial h}{\partial x^i},\qquad (i=1, \cdots, n). \tag{21}$$

记

$$\lambda_i(t)=\psi_0\frac{\partial h(t, x(t), y(t), \dot{x}(t), u(t))}{\partial \dot{x}^i}+\psi_i(t),$$

$$(i=1, \cdots, n), \tag{22}$$

那末根据(20)、(21)和(22)，推知(18)成为

$$\frac{d\psi_i}{dt}=-\psi_0\frac{\partial h(t, x(t), y(t), \dot{x}(t), u(t))}{\partial x^i}$$

$$-\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t, x(t), y(t), u(t))}{\partial x^i}\lambda_\alpha(t),$$

$$(i=1, 2, \cdots, n),$$

$$
\frac{d\psi_{n+j}}{dt}=-\psi_0\frac{\partial h(t,\ x(t),\ y(t),\ \dot{x}(t),\ u(t))}{\partial y^j}
$$

$$
-\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t,\ x(t),\ y(t),\ u(t))}{\partial y^j}\lambda_\alpha(t),
$$

$$
(j=1,\ 2,\ \cdots,\ m), \tag{23}
$$

而(19)成为

$$
\psi_{n+j}(t)+\sum_{\alpha=1}^{n}\frac{\partial f^\alpha(t,\ x(t),\ y(t),\ u(t))}{\partial u^j}\lambda_\alpha(t)
$$

$$
+\psi_0\frac{\partial h(t,\ x(t),\ y(t),\ \dot{x}(t),\ u(t))}{\partial u^j}=0,
$$

$$
(j=1,\ \cdots,\ m). \tag{24}
$$

置 $$\lambda_0=-\psi_0,$$

$$
F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)
$$

$$
=\lambda_0 h(t,\ x,\ y,\ \dot{x},\ \dot{y})+\sum_{\alpha=1}^{n}(\dot{x}^\alpha-f^\alpha(t,\ x,\ y,\ \dot{y}))\lambda_\alpha
$$

$$
=\lambda_0 h(t,\ x,\ y,\ \dot{x},\ \dot{y})+\langle\lambda,\ \dot{x}-f(t,\ x,\ y,\ \dot{y})\rangle.
$$

那末

$$
\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial \dot{x}^i}
$$

$$
=\lambda_0\frac{\partial h(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t))}{\partial \dot{x}^i}+\lambda_i(t)
$$

$$
=\psi_i(t),\qquad (i=1,\ 2,\ \cdots,\ n);
$$

$$
\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial x^i}
$$

$$
=\lambda_0\frac{\partial h(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t))}{\partial x^i}
$$

$$
-\sum_{\alpha=1}^{n}\lambda_\alpha(t)\frac{\partial f^\alpha(t,\ x(t),\ y(t),\ \dot{y}(t))}{\partial x^i},
$$

$$
(i=1,\ 2,\ \cdots,\ n);
$$

根据上式和(23)，得到

$$\frac{\partial F(t, x(t), y(t), \dot{x}(t), \dot{y}(t), \lambda_0, \lambda(t))}{\partial \dot{x}^i}$$
$$-\int_{t_0}^{t} \frac{\partial F(s, x(s), y(s), \dot{x}(s), \dot{y}(s), \lambda_0, \lambda(s))}{\partial x^i} ds = c_i$$
$$(i=1, 2, \cdots, n), \tag{25}$$

这里 c_i 是常数.

又根据(24)和(23), 成立着

$$\frac{\partial F(t, x(t), y(t), \dot{x}(t), \dot{y}(t), \lambda_0, \lambda(t))}{\partial \dot{y}^j}$$
$$= \lambda_0 \frac{\partial h(t, x(t), y(t), \dot{x}(t), \dot{y}(t))}{\partial \dot{y}^j}$$
$$-\sum_{\alpha=1}^{n} \lambda_\alpha(t) \frac{\partial f^\alpha(t, x(t), y(t), \dot{y}(t))}{\partial \dot{y}^j}$$
$$= \psi_{n+j}(t), \quad (j=1, 2, \cdots, m),$$
$$\frac{\partial F(t, x(t), y(t), \dot{x}(t), \dot{y}(t), \lambda_0, \lambda(t))}{\partial y^j}$$
$$= \lambda_0 \frac{\partial h(t, x(t), y(t), \dot{x}(t), \dot{y}(t))}{\partial y^j}$$
$$-\sum_{\alpha=1}^{n} \lambda_\alpha(t) \frac{\partial f^\alpha(t, x(t), y(t), \dot{y}(t))}{\partial y^j}$$
$$= \frac{d\psi_{n+j}(t)}{dt}, \quad (j=1, 2, \cdots, m).$$

所以得到

$$\frac{\partial F(t, x(t), y(t), \dot{x}(t), \dot{y}(t), \lambda_0, \lambda(t))}{\partial \dot{y}^j}$$
$$-\int_{t_0}^{t} \frac{\partial F(s, x(s), y(s), \dot{x}(s), \dot{y}(s), \lambda_0, \lambda(s))}{\partial y^j} ds = c_{n+j}$$
$$(j=1, 2, \cdots, m), \tag{26}$$

这里 c_{n+j} 是常数.

(25)和(26)写成微分方程组是

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right)=\frac{\partial F}{\partial x^i} \qquad (i=1,\ 2,\ \cdots,\ n)$$

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{y}^j}\right)=\frac{\partial F}{\partial y^j} \qquad (j=1,\ 2,\ \cdots,\ m). \tag{27}$$

它称为欧拉-拉格朗日方程组，而 F 称为拉格朗日函数．因为 $\psi_0^2+\|\psi(t)\|^2>0$，我们要推出 $\lambda_0^2+\|\lambda(t)\|^2>0$．否则，由 $\lambda_0=\lambda(t)=0$ 和(22)、(24)及 $\lambda_0=-\psi_0$ 推出 $\psi(t)=0$，它与 $\psi_0^2+\|\psi(t)\|^2>0$ 相矛盾．

由此，我们证得：如果 $x(\cdot)$ 和 $y(\cdot)$ 是拉格朗日问题的极值曲线，那末存在 λ_0 和 $\lambda(\cdot)$，使得

$$\lambda_0^2+\|\lambda(t)\|^2>0,$$

并且适合欧拉-拉格朗日方程组(27)．

对于 $t,\ x,\ y,\ \dot{x},\ \dot{y},\ \xi,\ v,\ \lambda_0,\ \lambda$，我们定义魏尔斯特拉斯(K. Weierstrass)函数

$$\begin{aligned}E&(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \xi,\ v,\ \lambda_0,\ \lambda)\\&=F(t,\ x,\ y,\ \xi,\ v,\ \lambda_0,\ \lambda)-F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)\\&\quad-\sum_{\alpha=1}^{n}(\xi^\alpha-\dot{x}^\alpha)\ \frac{\partial F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)}{\partial \dot{x}^\alpha}\\&\quad-\sum_{\beta=1}^{m}(v^\beta-\dot{y}^\beta)\ \frac{\partial F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)}{\partial \dot{y}^\beta}.\end{aligned} \tag{28}$$

设 $x=(\cdot)$，$y=y(\cdot)$ 是拉格朗日问题的极值曲线．设 $u(\cdot)=\dot{y}(\cdot)$，$\xi=f(t,\ x(t),\ y(t),\ v)$，$\lambda(\cdot)$ 由(22)决定，那末成立着

$$\frac{dx(t)}{dt}-f(t,\ x(t),\ y(t),\ u(t))=0,\quad \dot{y}(t)=u(t)$$

$$\xi-f(t,\ x(t),\ y(t),\ v)=0,$$

从而

$$\begin{aligned}F&(t,\ x(t),\ y(t),\ \dot{x}(t),\ u(t),\ \lambda_0,\ \lambda(t))\\&=\lambda_0 h(t,\ x(t),\ y(t),\ \dot{x}(t),\ u(t))\\&=-\psi_0 f^0(t,\ x(t),\ y(t),\ u(t));\end{aligned}$$

$$F(t,\ x(t),\ y(t),\ \xi,\ v,\ \lambda_0,\ \lambda(t))$$
$$=\lambda_0 h(t,\ x(t),\ y(t),\ \xi,\ v)=-\psi_0 f^0(t,\ x(t),\ y(t),\ v).$$

所以

$$E(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \xi,\ v,\ \lambda_0,\ \lambda(t))$$
$$=\psi_0 f^0(t,\ x(t),\ y(t),\ u(t))-\psi_0 f^0(t,\ x(t),\ y(t),\ v)$$
$$-\sum_{\alpha=1}^{n}(f^\alpha(t,\ x(t),\ y(t),\ v)$$
$$-f^\alpha(t,\ x(t),\ y(t),\ u(t)))\psi_\alpha(t)$$
$$-\sum_{\beta=1}^{m}(v^\beta-u^\beta(t))\psi_{n+\beta}(t)$$
$$=\mathscr{H}(t,\ \psi(t),\ x(t),\ y(t),\ u(t))$$
$$-\mathscr{H}(t,\ \psi(t),\ x(t),\ y(t),\ v).$$

根据最大值条件，得到

$$E(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \xi,\ v,\ \lambda_0,\ \lambda(t))\geqslant 0.$$

这就是极值曲线 $x=x(\cdot)$, $y=y(\cdot)$ 所应当满足的魏尔斯特拉斯不等式.

因此，利用最大值原理可以推出拉格朗日变分学问题的极值曲线所适合的一切必要条件(欧拉-拉格朗日方程组，魏尔斯特拉斯不等式). 即证明了下述定理.

定理 2 设 $x(\cdot)$, $y(\cdot)$ 是拉格朗日变分学问题的极值曲线，那末存在 λ_0 和 $\lambda(\cdot)$，使得 $\lambda_0^2+\|\lambda(t)\|^2>0$，并且适合

$$\begin{aligned}&\frac{d}{dt}\left(\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial \dot{x}}\right)\\&\quad=\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial x},\\&\frac{d}{dt}\left(\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial \dot{y}}\right)\\&\quad=\frac{\partial F(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \lambda_0,\ \lambda(t))}{\partial y},\end{aligned}\tag{29}$$

这里

$$F(t, x, y, \dot{x}, \dot{y}, \lambda_0, \lambda)$$
$$=\lambda_0 h(t, x, y, \dot{x}, \dot{y})+\langle\lambda, \dot{x}-f(t, x, y, \dot{y})\rangle. \tag{30}$$

同时，对于$\xi=f(t, x(t), y(t), v)$, $v\in\mathbb{R}^m$，成立着

$$E(t, x(t), y(t), \dot{x}(t), \dot{y}(t), \xi, v, \lambda_0, \lambda(t))\geqslant 0, \tag{31}$$

这里

$$E(t, x, y, \dot{x}, \dot{y}, \xi, v, \lambda_0, \lambda)$$
$$=F(t, x, y, \xi, v, \lambda_0, \lambda)-F(t, x, y, \dot{x}, \dot{y}, \lambda_0, \lambda)$$
$$-\left\langle\xi-\dot{x}, \frac{\partial F(t, x, y, \dot{x}, \dot{y}, \lambda_0, \lambda)}{\partial\dot{x}}\right\rangle$$
$$-\left\langle v-\dot{y}, \frac{\partial F(t, x, y, \dot{x}, \dot{y}, \lambda_0, \lambda)}{\partial\dot{y}}\right\rangle. \tag{32}$$

如果在拉格朗日问题中，边值条件(12)改为

$$x(t_0)=x_0, \quad x(t_1)=x_1,$$
$$y(t_0)=y_0.$$

那末在上述定理 2 的必要条件中，应补充横截条件

$$\frac{\partial F(t_1, x(t_1), y(t_1), \dot{x}(t_1), \dot{y}(t_1), \lambda_0, \lambda(t_1))}{\partial\dot{y}}=0. \tag{33}$$

三 从变分学的定理 2 推导最大值原理

现在用变分学来讨论§1的最优控制问题.

设系统的状态方程为

$$\frac{dx}{dt}=f(t, x, u(t)), \quad u(t)\in\mathbb{R}^m, \tag{34}$$

边值条件为

$$x(t_0)=x_0, \quad x(t_1)=x_1, \tag{35}$$

性能指标为

$$J(u(\cdot))=\int_{t_0}^{t_1} f^0(t, x(t), u(t))dt. \tag{36}$$

若记

$$y(t)=\int_{t_0}^{t}u(s)ds,\quad h(t,\ x,\ y,\ \dot{x},\ \dot{y})=f^0(t,\ x,\ \dot{y}),$$

那末
$$\frac{dy(t)}{dt}=u(t),\quad y(t_0)=0,$$

方程组(34)可写为

$$\frac{dx}{dt}-f\left(t,\ x,\ \frac{dy}{dt}\right)=0,\tag{37}$$

边值条件为

$$\begin{aligned}&x(t_0)=x_0,\ x(t_1)=0,\\&y(t_0)=0,\end{aligned}\tag{38}$$

泛函(36)成为

$$\int_{t_0}^{t_1}f^0(t,\ x(t),\ \dot{y}(t))dt.\tag{39}$$

这样,就把最优控制问题(34)、(35)、(36)转化为变分学的拉格朗日问题(37)、(38)、(39). 根据定理 2,如果置

$$\begin{aligned}&F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)\\&\qquad=\lambda_0f^0(t,\ x,\ \dot{y})+\langle\lambda,\ \dot{x}-f(t,\ x,\ \dot{y})\rangle,\end{aligned}$$

$$\begin{aligned}&E(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \xi,\ v,\ \lambda_0,\ \lambda)\\&\quad=F(t,\ x,\ y,\ \xi,\ v,\ \lambda_0,\ \lambda)-F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)\\&\qquad-\left\langle\xi-\dot{x},\ \frac{\partial F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)}{\partial\dot{x}}\right\rangle\\&\qquad-\left\langle v-\dot{y},\ \frac{\partial F(t,\ x,\ y,\ \dot{x},\ \dot{y},\ \lambda_0,\ \lambda)}{\partial\dot{y}}\right\rangle,\end{aligned}$$

那末存在 λ_0 和 $\lambda(\cdot)$ 满足

$$\lambda_0^2+\|\lambda(t)\|^2>0,$$

$$\frac{d\lambda(t)}{dt}=\lambda_0\frac{\partial f^0(t,\ x(t),\ \dot{y}(t))}{\partial x}-\frac{\partial f(t,\ x(t),\ \dot{y}(t))}{\partial x}\lambda(t),$$

$$\frac{d}{dt}\left\{\lambda_0\frac{\partial f^0(t,\ x(t),\ \dot{y}(t))}{\partial\dot{y}}-\frac{\partial f(t,\ x(t),\ \dot{y}(t))}{\partial\dot{y}}\lambda(t)\right\}=0,$$

$$\lambda_0\frac{\partial f^0(t_1,\ x(t_1),\ \dot{y}(t_1)}{\partial\dot{y}}-\frac{\partial f(t_1,\ x(t_1),\ \dot{y}(t_1))}{\partial\dot{y}}\lambda(t_1)=0.$$

设

$$\mathscr{H}(t,\ \psi_0,\ \psi,\ x,\ u)=\psi_0 f^0(t,\ x,\ u)+\langle\psi,\ f(t,\ x,\ u)\rangle,$$
$$\psi_0=-\lambda_0,\ \psi(\cdot)=\lambda(\cdot),\ u(\cdot)=\dot{y}(\cdot)$$

那末由上式得到

$$\frac{d\psi(t)}{dt}=-\psi_0\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\psi(t),$$

$$\frac{\partial\mathscr{H}(t,\ \psi_0,\ \psi(t),\ x(t),\ u(t))}{\partial u}$$
$$=\psi_0\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial u}+\frac{\partial f(t,\ x(t),\ u(t))}{\partial u}\psi(t)=0,$$

即
$$\frac{\partial\mathscr{H}(t,\ \psi_0,\psi(t),x(t),u(t))}{\partial u}=0.$$

又
$$\begin{aligned}0&\leqslant E(t,\ x(t),\ y(t),\ \dot{x}(t),\ \dot{y}(t),\ \xi,\ v,\ \lambda_0,\ \lambda(t))\\&=\lambda_0 f^0(t,\ x(t),\ v)-\lambda_0 f^0(t,\ x(t),\ u(t))\\&\quad-\langle f(t,\ x(t),\ v)-f(t,\ x(t),\ u(t)),\ \lambda(t)\rangle\\&\quad-\langle v-u(t),\ 0\rangle\\&=\mathscr{H}(t,\ \psi_0,\ \psi(t),\ x(t),\ u(t))\\&\quad-\mathscr{H}(t,\ \psi_0,\ \psi(t),\ x(t),\ v).\end{aligned}$$

所以

$$\mathscr{H}(t,\ \psi_0,\ \psi(t),\ x(t),\ u(t))=\max_{v\in\mathbb{R}^m}\mathscr{H}(t,\ \psi_0,\ \psi(t),\ x(t),\ v).$$ ∎

由此看出，在 $U=\mathbb{R}^m$ 时，最大值条件等价于魏尔斯特拉斯不等式，伴随方程相当于欧拉-拉格朗日方程.

必须指出，当 U 是任一集时，最大值条件是成立的，但是魏尔斯特拉斯不等式可能不成立. 所以，最优控制的最大值原理推广了古典变分学的拉格朗日问题的必要条件.

习　　题

1　试利用终端条件为 $x(t_1, v)\in Q$ 的最优控制问题的最大值原理讨论拉格朗日变分问题: 在约束条件

$$\frac{dx(t)}{dt}-f\left(t,\ x(t),\ y(t),\ \frac{dy(t)}{dt}\right)=0,$$
$$x(t_0)=x_0,\quad x(t_1)=x_1,$$
$$y(t_0)=y_0$$

下, 求泛函

$$\int_{t_0}^{t_1} h\left(t,\ x(t),\ y(t),\ \frac{dx(t)}{dt},\ \frac{dy(t)}{dt}\right)dt$$

的极小值.

§4　线性系统的二次最优控制问题

设系统的状态方程为

$$\begin{aligned}&\frac{dx(t)}{dt}=A(t)x(t)+B(t)u(t),\\ &x(t_0)=x_0,\end{aligned}\tag{1}$$

其中 $A(\cdot)$: $[t_0,\ t_1]\to\mathbb{R}^{n\times n}$ 是可积的, $B(\cdot)$: $[t_0,\ t_1]\to\mathbb{R}^{n\times m}$ 是平方可积的, $u(\cdot)$: $[t_0,\ t_1]\to\mathbb{R}^m$ 是平方可积的, $x_0\in\mathbb{R}^n$ 是给定的.

根据第一章§5, 对于 $u(\cdot)\in L^2([t_0,\ t_1];\ \mathbb{R}^m)$, 方程组(1)以 $(t_0,\ x_0)$ 为初值的解 $x(\cdot)=x(\cdot,\ u)$ 在 $[t_0,\ t_1]$ 上存在.

又设 $Q(\cdot)$: $[t,\ t_1]\to\mathbb{R}^{n\times n}$ 和 $R(\cdot)$: $[t_0,\ t_1]\to\mathbb{R}^{m\times m}$ 是连续的, 并且

$$Q^{T}(t)\equiv Q(t)\geqslant 0,$$
$$R^{T}(t)\equiv R(t)>0,$$

即 Q 是非负定的, R 是正定的. 再设 $Q_1\in\mathbb{R}^{n\times n}$ 是非负定的对称矩阵.

对于 $u(\cdot)\in L^2([t_0,\ t_1];\ \mathbb{R}^m)$, 考虑二次泛函

$$J(u(\cdot))=\frac{1}{2}\langle Q_1x(t_1),\ x(t_1)\rangle$$
$$+\frac{1}{2}\int_{t_0}^{t_1}\{\langle Q(t)x(t),\ x(t)\rangle+\langle R(t)u(t),\ u(t)\rangle\}dt. \tag{2}$$

线性系统的二次最优控制问题是：寻求平方可积的 $u(\cdot)$ 使不等式

$$J(u(\cdot))\leqslant J(v(\cdot))$$

对于 $v(\cdot)\in L^2([t_0,\ t_1];\ \mathbb{R}^m)$ 和(1)、(2)皆成立。

定理 1 线性系统的二次最优控制问题的解 $u(\cdot)$ 是存在唯一的。

它的证明类似定理 II-7-1，请读者完成。

定理 2 设 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,\ u)$ 是线性系统二次最优控制问题的解，那末存在 $\psi(\cdot)$ 适合

$$\begin{aligned}&\frac{d\psi(t)}{dt}=-A^T(t)\psi(t)+Q(t)x(t),\\&\psi(t_1)=-Q_1x(t_1),\end{aligned} \tag{3}$$

使得最优控制 $u(\cdot)$ 满足关系

$$u(t)=R^{-1}(t)B^T(t)\psi(t). \tag{4}$$

证 读者可以用第二章 § 7 的方法进行。现在，利用最大值原理来证明它。

首先，形如(2)的泛函不是 § 1 所讨论的性能指标，因此我们利用

$$\begin{aligned}\langle Q_1x(t_1),\ x(t_1)\rangle&=\int_{t_0}^{t_1}\frac{d}{dt}\langle Q_1x(t),\ x(t)\rangle dt+\langle Q_1x_0,\ x_0\rangle\\&=\langle Q_1x_0,\ x_0\rangle\\&\quad+\int_{t_0}^{t_1}\langle (A^T(t)Q_1+Q_1A(t))x(t),\ x(t)\rangle dt\\&\quad+2\int_{t_0}^{t_1}\langle B^T(t)Q_1x(t),\ u(t)\rangle dt,\end{aligned}$$

把(2)化为

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,\ x(t),\ u(t))dt+\langle Q_1x_0,\ x_0\rangle/2,$$

其中

$$f^0(t,\ x,\ u)=\frac{1}{2}\langle[Q(t)+A^T(t)Q_1+Q_1A(t)]x,\ x\rangle$$
$$+\langle B^T(t)Q_1x,\ u\rangle+\frac{1}{2}\langle R(t)u,\ u\rangle.$$

根据最大值原理，存在 $p_0\leqslant 0$ 和 $p(\cdot)$ 适合

$$p_0^2+\|p(t)\|^2>0,$$

$$\frac{dp(t)}{dt}=-\frac{\partial\mathscr{H}(t,\ p_0,\ p(t),\ x(t),\ u(t))}{\partial x}$$
$$=-p_0\{(Q(t)+A^T(t)Q_1+Q_1A(t))x(t)+Q_1B(t)u(t)\}$$
$$-A^T(t)p(t),\tag{5}$$

$$p(t_1)=0,\tag{6}$$

$$0=\frac{\partial\mathscr{H}(t,\ p_0,\ p(t),\ x(t),\ u(t))}{\partial u}$$
$$=B^T(t)p(t)+\{B^T(t)Q_1x(t)+Ru(t)\}p_0\tag{7}$$

如果 $p_0=0$，那末由(5)和(6)得到 $p(t)=0$，它与 $p_0^2+\|p(t)\|^2>0$ 矛盾，所以 $p_0\neq 0$，取 $p_0=-1$.

由(7)得到

$$B^T(t)(p(t)-Q_1x(t))-Ru(t)=0,$$

而方程组(5)成为

$$\frac{dp(t)}{dt}=(Q(t)+A^T(t)Q_1+Q_1A(t))x(t)+Q_1B(t)u(t)$$
$$-A^T(t)p(t).$$

记 $\psi(\cdot)=p(\cdot)-Q_1x(\cdot)$，那末

$$B^T(t)\psi(t)-Ru(t)=0,\tag{8}$$

$$\frac{d\psi(t)}{dt}=\frac{dp(t)}{dt}-Q_1\frac{dx(t)}{dt}=Q(t)x(t)-A^T(t)\psi(t),$$

$$\psi(t_1) = -Q_1 x(t_1).$$

它就是(4)和(3). ∎

定理3 设 $P(\cdot):[t_0, t_1]\to\mathbb{R}^{n\times n}$ 满足黎卡提方程

$$\begin{aligned}
-\frac{dP(t)}{dt} &= Q(t)+A^T(t)P(t)+P(t)A(t) \\
&\quad -P(t)B(t)R^{-1}(t)B^T(t)P(t),
\end{aligned} \tag{9}$$

$$P(t_1)=Q_1,$$

那末线性系统二次最优控制问题的解 $u(\cdot)$ 和 $x(\cdot)=x(\cdot, u)$ 适合关系式

$$u(t)\equiv -R^{-1}(t)B^T(t)P(t)x(t). \tag{10}$$

证 因为最优控制的解 $u(\cdot)$ 适合关系式(4),所以

$$u(t)\equiv R^{-1}(t)B^T(t)\psi(t).$$

而 $\psi(\cdot)$ 和 $x(\cdot)$ 适合(3)和(1),所以

$$\frac{dx(t)}{dt}\equiv A(t)x(t)+B(t)R^{-1}(t)B^T(t)\psi(t),$$

$$\frac{d\psi(t)}{dt}\equiv Q(t)x(t)-A^T(t)\psi(t),$$

$$\psi(t_1)=-Q_1x(t_1), \quad x(t_0)=x_0.$$

从而

$$\begin{aligned}
&\frac{d}{dt}\{\psi(t)+P(t)x(t)\} \\
&\equiv Q(t)x(t)-A^T(t)\psi(t)+\frac{dP(t)}{dt}x(t)+P(t)A(t)x(t) \\
&\quad +P(t)B(t)R^{-1}(t)B^T(t)\psi(t).
\end{aligned}$$

利用(9)和上式,得到

$$\begin{aligned}
&\frac{d}{dt}\{\psi(t)+P(t)x(t)\} \\
&\equiv\{-A^T(t)+P(t)B(t)R^{-1}(t)B^T(t)\}(\psi(t)+P(t)x(t)),
\end{aligned}$$

但是 $$\psi(t_1)+P(t_1)x(t_1)=0,$$

所以,从微分方程解对初值的唯一性即得知

$$\psi(t)+P(t)x(t)\equiv 0$$

在 $[t_0, t_1]$ 上成立，即 $\psi(t)\equiv -P(t)x(t)$. 把它代入(4)就得到(10).

等式(10)表明：最优控制是状态的线性函数．我们称它为最优线性反馈调节器。

习　题

1　证明定理1.

2　利用第二章§7的方法证明定理2.

3　利用第二章§7的方法证明定理3.

第四章　最大值原理的证明

§1　引　　言

最大值原理在最优控制理论、变分学中的重要地位，已从第三章看出了．在那里，我们就线性系统的时间最优控制问题、变分学中拉格朗日问题以及线性系统二次最优控制问题的必要条件的推导，指出从最大值原理可以推出上述种种问题的一切必要条件．

本章给出最大值原理的严格证明．

设受控系统的状态方程为

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ u(t)), \tag{1}$$

其中 $f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 满足下述条件：

1° 对于 $(x,\ u)\in\mathbb{R}^n\times R^m$，$f(\cdot,\ x,\ u):[t_0,\ t_1]\to\mathbb{R}^n$ 是可积的；

2° 对于 $(t,\ u)\in\mathbb{R}\times\mathbb{R}^m$，$f(t,\ \cdot,\ u)$：$\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微的；

3° 对于 $(t,\ x)\in\mathbb{R}\times\mathbb{R}^n$，$f(t,\ x,\ \cdot):\mathbb{R}^m\to\mathbb{R}^n$ 是连续的．

设 $x=(x^1,\ x^2,\ \cdots,\ x^n)^T\in\mathbb{R}^n$，$u=(u^1,\ u^2,\ \cdots,\ u^m)^T\in\mathbb{R}^m$，而 $f=(f^1,\ f^2,\ \cdots,\ f^n)^T\in\mathbb{R}^n$，那末 $\frac{\partial f}{\partial x}$ 是下面矩阵的记号：

$$\begin{bmatrix}\frac{\partial f^1}{\partial x^1} & \frac{\partial f^2}{\partial x^1} & \cdots & \frac{\partial f^n}{\partial x^1}\\ \frac{\partial f^1}{\partial x^2} & \frac{\partial f^2}{\partial x^2} & \cdots & \frac{\partial f^n}{\partial x^2}\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\partial f^1}{\partial x^n} & \frac{\partial f^2}{\partial x^n} & \cdots & \frac{\partial f^n}{\partial x^n}\end{bmatrix}.$$

又设 U 是 $\mathbb{R}^m$ 中给定的集，它表征对控制器的限制，称为系统的控制区域. 如果 $u(\cdot):[t_0,\ t_1]\to\mathbb{R}^m$ 是有界可积的，并且在 $[t_0, t_1]$ 上几乎处处成立着 $u(t)\in U$，就称 $u(\cdot)$ 在 $[t_0,\ t_1]$ 上是系统(1)的容许控制. 记为

$$u(\cdot)\in U_{ad}.$$

设 $(t_0,\ x_0)\in\mathbb{R}\times\mathbb{R}^n$ 和 $(t_1,\ x_1)\in\mathbb{R}\times\mathbb{R}^n$ 给定.

对于 $u(\cdot)\in U_{ad}$，方程组(1)以 (t_0, x_0) 为初值的解记为 $x(\cdot)=x(\cdot,\ u)$，即

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ u(t)),$$

$$x(t_0)=x_0,$$

或
$$x(t)=x_0+\int_{t_0}^{t}f(s,\ x(s),\ u(s))ds.$$

如果 $x(\cdot)$ 在 $[t_0,\ t_1]$ 上有意义，并且

$$x(t_1)=x_1, \tag{2}$$

我们就称控制作用 $u(\cdot)$ 把系统(1)从状态 x_0（或初值 $(t_0,\ x_0)$）迁移到状态 x_1（或终值 $(t_1,\ x_1)$）.

再设 f^0: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 满足 f 类似的条件 1°、2° 和 3°，这里

$$\frac{\partial f^0}{\partial x}=\left(\frac{\partial f^0}{\partial x^1}\ \frac{\partial f^0}{\partial x^2}\cdots\frac{\partial f^0}{\partial x^n}\right)^T.$$

对于把系统(1)从状态 x_0 迁移到状态 x_1 的容许控制 $u(\cdot)$，考虑积分

$$\int_{t_0}^{t_1}f^0(t,\ x(t),\ u(t))dt, \tag{3}$$

它是容许控制 $u(\cdot)$ 的泛函，称为系统的性能指标，记为 $J(u(\cdot))$.

最优控制问题 选取把系统(1)从状态 x_0 迁移到状态 x_1 的容许控制 $u(\cdot)\in U_{ad}$，使得

$$J(u(\cdot))$$
$$=\inf\{J(v(\cdot))\,|\,x(t_0,\ v)=x_0,\ x(t_1,\ v)=x_1,\ v(\cdot)\in U_{ad}\}. \tag{4}$$

这里假定$(t_0,\ x_0)$和$(t_1,\ x_1)$是给定的.

为使所讨论的最优控制问题能包含时间最优控制问题，还需讨论 t_1 不固定的控制问题，即假设$(t_0,\ x_0)\in\mathbb{R}\times\mathbb{R}^n$ 和 $x_1\in\mathbb{R}^n$ 给定，而要求 t_1 是使得(2)成立的时刻. 这时，对于不同的容许控制 $v(\cdot)\in U_{ad}$，使(2)成立的时刻 t_1 可能是不同的.

当 $f^0\equiv1$，t_1 不固定时，$J(u(\cdot))=t_1-t_0$ 成为系统(1)从状态 x^0 到状态 x_1 所需的时间，问题就是时间最优控制问题.

如果容许控制 $u(\cdot)$ 把系统(1)从状态 x_0 迁移到状态 x_1，并且满足关系式(4)，就称 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,\ u)$ 为最优控制问题的解，称 $u(\cdot)$ 为最优控制，$x(\cdot)$ 为最优轨线. 最优控制理论的重要内容之一是：给出最优控制 $u(\cdot)$ 和最优轨线 $x(\cdot)$ 所适合的条件. 最大值原理正是回答这一问题的.

最大值原理 设 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,\ u)$ 是最优控制问题(4)的解，那末存在 $\psi_0\leqslant0$ 和 $\psi(\cdot)$ 满足

$$\psi_0^2+\|\psi(t)\|^2>0,$$
$$\frac{d\psi}{dt}=-\psi_0\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\psi(t), \tag{5}$$

并且最大值条件

$$\psi_0f^0(t,\ x(t),\ u(t))+\langle\psi(t),\ f(t,\ x(t),\ u(t))\rangle$$
$$=\max^{v\in U}\{\psi_0f^0(t,\ x(t),\ v)+\langle\psi(t),\ f(t,\ x(t),\ v)\rangle\} \tag{6}$$

在$[t_0,\ t_1]$上几乎处理成立.

$\mathscr{H}:\mathbb{R}\times\mathbb{R}^{n+1}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 的定义是

$$\mathscr{H}(t, \psi, x, u)=\psi_0 f^0(t, x, u)+\langle\psi, f(t, x, u)\rangle,$$

称为最优控制问题的哈密顿函数.

这时,方程组(1)和(5)可写为

$$\frac{dx(t)}{d}=\frac{\partial\mathscr{H}(t, \psi(t), x(t), u(t))}{\partial\psi},$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(t, \psi(t), x(t), u(t))}{\partial x}.$$

而最大值条件(6)成为

$$\mathscr{H}(t, \psi(t), x(t), u(t))$$
$$=\max_{v\in U}\mathscr{H}(t, \psi(t), x(t), v).$$

如果补充引进 $x^0(\cdot)$,适合

$$\frac{dx^0(t)}{dt}=f^0(t, x(t), u(t)),\ x^0(t_0)=0,$$

那末
$$x^0(t_1)=\int_{t_0}^{t_1}f^0(t, x(t), u(t))dt.$$

若记

$$\boldsymbol{x}=\begin{pmatrix}x^0\\ x\end{pmatrix},\quad \boldsymbol{f}=\begin{pmatrix}f^0\\ f\end{pmatrix},\quad \boldsymbol{\psi}=\begin{pmatrix}\psi_0\\ \psi\end{pmatrix},$$

那末

$$\frac{d\boldsymbol{x}(t)}{dt}=\boldsymbol{f}(t, x(t), u(t)),\tag{7}$$

$$\boldsymbol{x}(t_0)=\boldsymbol{x}_0=\begin{pmatrix}0\\ x_0\end{pmatrix},$$

$$\boldsymbol{x}(t_1)=\begin{pmatrix}J(u(\cdot))\\ x_1\end{pmatrix}.$$

因此,最优控制问题等价于要求:方程组(7)以$(t_0, \boldsymbol{x}_0)$为初值的解,满足

$$x(t_1)=x_1,$$

并且 $\boldsymbol{x}(t_1)$ 的第 0 坐标 $x^0(t_1)$ 取最小值.

本章还将讨论最优控制问题的种种推广。

§2 轨线的变分

设 $u(\cdot)\in U_{ad}$，系统

$$\frac{dx}{dt}=f(t,\ x,\ u(t)) \tag{1}$$

以 $(t_0,\ x_0)$ 为初值的解 $x=x(\cdot)=x(\cdot,\ u)$ 在 $[t_0,\ t_1]$ 上存在。

对于另一 $v(\cdot)\in U_{ad}$，我们不能肯定方程组(1)以 $(t_0,\ x_0)$ 为初值的解 $x(\cdot,\ v)$ 也在 $[t_0,\ t_1]$ 上存在。但是，我们要设法根据 $v(\cdot)$ 和 $u(\cdot)$，来构造容许控制 $u^\varepsilon(\cdot)$，使得它所相应的解 $x(\cdot,\ u^\varepsilon)$ 在 $[t_0,\ t_1]$ 上存在。

为此，我们考察

$$f(\cdot,\ x(\cdot),\ v(\cdot))-f(\cdot,\ x(\cdot),\ u(\cdot)),$$

它在 $[t_0,\ t_1]$ 上是可积的。

根据定理 I-1-4，对于 $\varepsilon>0$，存在可测集 $E_\varepsilon\subset[t_0,\ t_1]$，使得

$$\mu(E_\varepsilon)=\varepsilon(t_1-t_0),$$

$$\begin{aligned}\varepsilon\int_{t_0}^{t}\{&f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds\\ =&\int_{E_\varepsilon\cap[t_0,\ t]}\{f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds\\ &+\gamma(t),\end{aligned} \tag{2}$$

$$\|\gamma(t)\|<\varepsilon^2.$$

置

$$u^\varepsilon(t)=\begin{cases}v(t), & t\in E_\varepsilon,\\ u(t), & t\in[t_0,\ t_1]\setminus E_\varepsilon,\end{cases}$$

那末 $u^\varepsilon(\cdot)\in U_{ad}$，(2)可写为

$$\varepsilon\int_{t_0}^{t}\{f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds$$

$$=\int_{t_0}^{t}\{\boldsymbol{f}(s, x(s), u^{\varepsilon}(s))-\boldsymbol{f}(s, x(s), u(s))\}ds+\boldsymbol{\gamma}(t).$$

(3)

设 $\boldsymbol{x}^{\varepsilon}(\cdot)=\boldsymbol{x}(\cdot, u^{\varepsilon})$ 适合

$$\boldsymbol{x}^{\varepsilon}(t)=\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{f}(s, \boldsymbol{x}^{\varepsilon}(s), u^{\varepsilon}(s))ds,$$

那末因为对于 $t\in[t_0, t_1]$ 和 $\boldsymbol{x}\in\mathbb{R}^{n+1}$ 成立着

$$\lim_{\varepsilon\to 0}\int_{t_0}\boldsymbol{f}(s, x, u^{\varepsilon}(s))ds=\int_{t_0}^{t}\boldsymbol{f}(s, x, u(s))ds,$$

根据定理I-4-1，当 $\varepsilon>0$ 充分小时，$\boldsymbol{x}^{\varepsilon}(\cdot)$ 在 $[t_0, t_1]$ 上存在，并且

$$\lim_{\varepsilon\to 0}\boldsymbol{x}^{\varepsilon}(t)=\boldsymbol{x}(t) \tag{4}$$

在 $[t_0, t_1]$ 上一致地成立.

进而，我们有下述定理.

定理 1 成立着关系式

$$\boldsymbol{x}^{\varepsilon}(t)=\boldsymbol{x}(t)+\varepsilon\delta\boldsymbol{x}(t)+\boldsymbol{\eta}(t), \tag{5}$$

其中

$$\|\boldsymbol{\eta}(t)\|=o(\varepsilon),$$

而 $\delta\boldsymbol{x}(\cdot)$ 适合变分方程组

$$\frac{d}{dt}\delta\boldsymbol{x}(t)=\left(\frac{\partial\boldsymbol{f}(t, x(t), u(t))}{\partial x}\right)^{T}\delta x(t)$$

$$+\boldsymbol{f}(t, x(t), v(t))-\boldsymbol{f}(t, x(t), u(t)), \tag{6}$$

$$\delta\boldsymbol{x}(t_0)=\mathbf{0}.$$

证 根据 $\boldsymbol{x}^{\varepsilon}(\cdot)$ 和 $\boldsymbol{x}(\cdot)$ 适合方程组

$$\boldsymbol{x}^{\varepsilon}(t)=\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{f}(s, \boldsymbol{x}^{\varepsilon}(s), u^{\varepsilon}(s))ds,$$

$$\boldsymbol{x}(t)=\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{f}(s, x(s), u(s))ds,$$

所以 $$\boldsymbol{y}^{\varepsilon}(\cdot)=\frac{1}{\varepsilon}\{\boldsymbol{x}^{\varepsilon}(\cdot)-\boldsymbol{x}(\cdot)\}$$

适合

$$y^{\varepsilon}(t)=\frac{1}{\varepsilon}\int_{t_0}^{t}\{f(s,\ x^{\varepsilon}(s),\ u^{\varepsilon}(s))-f(s,\ x(s),\ u^{\varepsilon}(s))\}ds$$

$$+\frac{1}{\varepsilon}\int_{t_0}^{t}\{f(s,\ x(s),\ u^{\varepsilon}(s))-f(s,\ x(s),\ u(s))\}ds$$

$$=\int_{t_0}^{t}\int_{0}^{1}\sum_{j=1}^{n}\frac{\partial f(s,\ x(s)+\tau(x^{\varepsilon}(s)-x(s)),\ u^{\varepsilon}(s))}{\partial x^{j}}y^{j\varepsilon}(s)d\tau ds$$

$$+\int_{t_0}^{t}\{f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds$$

$$+\frac{1}{\varepsilon}\gamma(t). \tag{7}$$

当 $t\in[t_0,\ t_1]$，$y\in\mathbb{R}^{n+1}$ 时，成立着

$$\int_{t_0}^{t}\sum_{j=1}^{n}\int_{0}^{1}\frac{\partial f(s,\ x(s)+\tau(x^{\varepsilon}(s)-x(s)),\ u^{\varepsilon}(s))}{\partial x^{j}}d\tau y^{j}ds$$

$$=\int_{t_0}^{t}\sum_{j=1}^{n}\int_{0}^{1}\frac{\partial f(s,\ x(s)+\tau(x^{\varepsilon}(s)-x(s)),\ u(s))}{\partial x^{j}}d\tau y^{j}ds$$

$$-\int_{E_{\varepsilon}\cap[t_0,t]}\sum_{j=1}^{n}\int_{0}^{1}\frac{\partial f(s,\ x(s)+\tau(x^{\varepsilon}(s)-x(s)),\ u(s))}{\partial x^{j}}d\tau y^{j}ds$$

$$+\int_{E_{\varepsilon}\cap[t_0,t]}\sum_{j=1}^{n}\int_{0}^{1}\frac{\partial f(s,\ x(s)+\tau(x^{\varepsilon}(s)-x(s)),\ v(s))}{\partial x^{j}}d\tau y^{j}ds.$$

当 $\varepsilon\to 0$，上式的极限是

$$\int_{t_0}^{t}\sum_{j=1}^{n}\frac{\partial f(s,\ x(s),\ u(s))}{\partial x^{j}}y^{j}ds$$

$$=\int_{t_0}^{t}\left(\frac{\partial f(s,\ x(s),\ u(s))}{\partial x}\right)^{T}y\,ds.$$

根据定理 I–4–1，由关系式(7)推知：极限

$$\lim_{\varepsilon\to 0}y^{\varepsilon}(\cdot)=\delta x(\cdot)$$

存在，并且 $\delta x(\cdot)$ 适合

$$\delta x(t)=\int_{t_0}^{t}\left(\frac{\partial f(s,\ x(s),\ u(s))}{\partial x}\right)^{T}\delta x(s)ds$$

$$+\int_{t_0}^{t}\{f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds.$$

即 $\delta x(\cdot)$ 适合关系(6).

因此 $$x^{\varepsilon}(\cdot)=x(\cdot)+\varepsilon\delta x(\cdot)+o(\varepsilon).$$ ∎

记

$$W=\{x\,|\,x=\varepsilon\delta x(t_1),\ \varepsilon\geqslant 0,\ v(\cdot)\in U_{ad}\}.$$

定理 2 W 是 $\mathbb{R}^{n+1}$ 中的凸锥.

证 根据定理 II-6-1,集

$$R(t_1)=\{x\,|\,x=\delta x(t_1),\ v(\cdot)\in U_{ad}\}$$

是凸集. 从而 W 是凸锥 ∎

习 题

1 对于系统(1),设 $x=x(\cdot,\ u)$ 是对应于 $u(\cdot)\in U_{ad}$ 的以 $(t_0,\ x_0)$ 为初值的解. 又设 $\lambda_1\geqslant 0,\ \lambda_2\geqslant 0,\ \cdots,\ \lambda_m\geqslant 0,\ \lambda_1+\lambda_2+\cdots+\lambda_m\leqslant 1,\ v_j(\cdot)\in U_{ad}\ (j=1,\ 2,\ \cdots,\ m)$. 试证: 存在容许控制 $u^{\varepsilon}(\cdot,\ \lambda)\in U_{ad}$,使得系统(1)相应于 $u^{\varepsilon}(\cdot,\ \lambda)$ 的以 $(t_0,\ x_0)$ 为初值的解 $x=x_\lambda^{\varepsilon}(\cdot)=x(\cdot,\ u^{\varepsilon}(\cdot,\ \lambda))$ 满足

$$x_\lambda^{\varepsilon}(\cdot)=x(\cdot)+\varepsilon\delta x_\lambda(\cdot)+o(\varepsilon),$$

这里 $\delta x_\lambda(\cdot)$ 适合

$$\delta x_\lambda(t)=\int_{t_0}^{t}\left(\frac{\partial f(s,\ x(s),\ u(s))}{\partial x}\right)^T\delta x_\lambda(s)ds$$
$$+\sum_{j=1}^{m}\lambda_j\int_{t_0}^{t}\{f(s,\ x(s),\ v_j(s))-f(s,\ x(s),\ u(s))\}ds.$$

并且 $x_\lambda^{\varepsilon}(\cdot)$ 关于 λ 是连续的.

§3 最大值原理的证明

一 终端自由的情形

首先讨论

$$x(t_1)\in\mathbb{R}^n,$$

即终端无约束的情形.

根据定理 2-1，我们有

$$J(u^{\varepsilon}(\cdot))=J(u(\cdot))+\varepsilon\int_{t_{0}}^{t_1}\left\langle\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x},\ \ \delta x(t)\right\rangle dt$$

$$+\varepsilon\int_{t_0}^{t_1}\{f^0(t,\ x(t),\ v(t))-f^0(t,\ x(t),\ u(t))\}dt$$

$$+o(\varepsilon). \tag{1}$$

因为 $u(\cdot)$ 是最优的，$x(t_1)$ 自由，所以

$$J(u^{\varepsilon}(\cdot))\geqslant J(u(\cdot)),$$

从而由(1)得到

$$\int_{t_0}^{t_1}\left\langle\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x},\ \delta x(t)\right\rangle dt$$

$$+\int_{t_0}^{t_1}\{f^0(s,\ x(s),\ v(s))-f^0(s,\ x(s),\ u(s))\}ds\geqslant 0. \tag{2}$$

设 $\tilde{\psi}(\cdot)$ 适合

$$\frac{d\tilde{\psi}(t)}{dt}=\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\ \tilde{\psi}(t), \tag{3}$$

$$\tilde{\psi}(t_1)=0,$$

而 $\delta x(\cdot)$ 适合

$$\frac{d\delta x(t)}{dt}=\left(\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\right)^T\delta x(t)$$

$$+f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t)), \tag{4}$$

$$\delta x(t_0)=0,$$

根据第一章 § 5 公式(13)，推知(2)的第一项等于

$$-\int_{t_0}^{t_1}\langle\tilde{\psi}(t),\ f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t))\rangle dt.$$

所以由(2)得到

$$\int_{t_0}^{t_1}\{-f^0(t,\ x(t),\ u(t))+\langle\tilde{\psi}(t),\ f(t,\ x(t),\ u(t))\rangle\}dt$$

$$\geqslant \int_{t_0}^{t_1} \{-f^0(t, x(t), v(t)) + \langle \tilde{\psi}(t), f(t, x(t), v(t)) \rangle\} dt. \tag{5}$$

现从(5)来推导最大值条件．设 $u \in U$ 是给定的，t 是函数

$$-f^0(\cdot, x(\cdot), u(\cdot)) + \langle \tilde{\psi}(\cdot), f(\cdot, x(\cdot), u(\cdot)) \rangle$$

和函数 $\quad -f^0(\cdot, x(\cdot), u) + \langle \tilde{\psi}(\cdot), f(\cdot, x(\cdot), u) \rangle$

的勒贝格点，$h>0$，取 $v(\cdot) \in U_{ad}$ 适合

$$v(s) = \begin{cases} u, & |s-t| \leqslant h, \\ u(s), & \text{其余}, \end{cases}$$

那末由(5)得到

$$\int_{t-h}^{t+h} \{-f^0(s, x(s), u(s)) + \langle \tilde{\psi}(s), f(s, x(s), u(s)) \rangle\} ds$$
$$\geqslant \int_{t-h}^{t+h} \{-f^0(s, x(s), u) + \langle \tilde{\psi}(s), f(s, x(s), u) \rangle\} ds.$$

在上式两端除以 $2h$，并令 $h \to 0$，得到

$$-f^0(t, x(t), u(t)) + \langle \tilde{\psi}(t), f(t, x(t), u(t)) \rangle$$
$$\geqslant -f^0(t, x(t), u) + \langle \tilde{\psi}(t), f(t, x(t), u) \rangle.$$

由于勒贝格点在 $[t_0, t_1]$ 是满测度的，所以上式在 $[t_0, t_1]$ 上是几乎处处成立的．在 U 中取可数多个点 $\{u_k\}$，它们在 U 中稠密．从而上式对于 $u_k \in U(k=1, 2, \cdots)$ 在 $[t_0, t_1]$ 上几乎处处成立．

因此，最大值条件

$$-f^0(t, x(t), u(t)) + \langle \tilde{\psi}(t), f(t, x(t), u(t)) \rangle$$
$$= \max_{v \in U} \{-f^0(t, x(t), v) + \langle \tilde{\psi}(t), f(t, x(t), v) \rangle\}$$

在 $[t_0, t_1]$ 上几乎处处成立． ▮

我们指出：在 $x(t_1) \in \mathbb{R}^n$ 时，

$$\psi_0 = -1,$$
$$\tilde{\psi}(t_1) = 0,$$

其中 $\psi_0 = -1$ 称为正则性条件；$\tilde{\psi}(t_1) = 0$ 称为横截条件．

二　最优值函数具李普希茨条件的情形

为了强调最优控制问题对终端状态 x_1 的依赖关系，我们记

$$V(x_1)=\inf\{J(v(\cdot))\,|\,x(t_0,\ v)=x_0,$$
$$x(t_1,\ v)=x_1,\ v(\cdot)\in U_{ad}\},$$

它表明泛函的最小值是终端状态 x_1 的函数，我们称 $V(\cdot)$ 为最优控制问题的值函数．这里，我们视 x_0 为固定的，x_1 是给定的．对不同的问题，x_1 是不同的．

在第三章讨论动态规划方程时，是假定 x_1 固定，考虑最优控制问题的最小值对初始状态的依赖关系．并且，在值函数 S 二阶连续可微时推导了最大值原理．但是，一般说来 S 不具有连续偏导数．

现在假设 $V(\cdot)$ 是满足局部李普希茨(R. Lipschitz)条件，即存在 $M>0$ 和 $\delta>0$，使得：当 $\|\bar{x}_1-x_1\|<\delta$ 时

$$|V(\bar{x}_1)-V(x_1)|\leqslant M\|\bar{x}_1-x_1\|. \tag{6}$$

我们将在条件(6)成立时来证明最大值原理．

设 $\Phi(\cdot,\ \cdot)$ 是齐次微分方程组

$$\frac{dy(t)}{dt}=\left(\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\right)^T y(t)$$

的转移矩阵，那末(4)的解 $\delta x(\cdot)$ 可表为

$$\delta x(t)=\int_{t_0}^{t}\Phi(t,\ s)\{f(s,\ x(s),\ v(s))-f(s,\ x(s),\ u(s))\}ds.$$

如果容许控制 $v(\cdot)$ 使得

$$\delta x(t_1)=0,$$

即满足

$$\int_{t_0}^{t_1}\Phi(t_1,\ t)\{f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t))\}dt=0, \tag{7}$$

那末由于　　$x^\varepsilon(t_1)=x_1+o(\varepsilon),$

和(6)推知

$$J(u^{\varepsilon}(\cdot))=J(u(\cdot))+\varepsilon\int_{t_0}^{t_1}\left\langle\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x},\ \delta x(t)\right\rangle dt$$
$$+\varepsilon\int_{t_0}^{t_1}\{f^0(t,\ x(t),\ v(t))-f^0(t,\ x(t),\ u(t))\}dt$$
$$+o(\varepsilon)$$
$$\geqslant V(x_1+o(\varepsilon))\geqslant V(x_1)+o(\varepsilon)$$
$$=J(u(\cdot))+o(\varepsilon).$$

由此得到不等式(2), 从而得到不等式(5).

设

$$b^0(t,\ v)=f^0(t,\ x(t),\ v)-f^0(t,\ x(t),\ u(t))$$
$$+\langle\tilde{\psi}(t),\ f(t,\ x(t),\ u(t))-f(t,\ x(t),\ v)\rangle$$

那末我们证得了下述定理.

定理 1 设 $u(\cdot)$ 和 $x(\cdot)$ 是最优控制问题的解, 并且值函数 $V(\cdot)$ 是局部李普希茨的, 那末 $u(\cdot)$ 也是下述线性最优控制问题的解, 即在线性状态方程

$$\frac{dz(t)}{dt}=\Phi(t_1,\ t)\{f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t))\},$$
$$z(t_0)=0,\quad z(t_1)=0,$$
$$v(\cdot)\in U_{ad}$$

下, 它使泛函

$$J_1(v(\cdot))=\int_{t_0}^{t_1}b^0(t,\ v(t))dt,$$

达到最小值 0, 并且

$$J_1(u(\cdot))=\inf\{J_1(v(\cdot))\,|\,v(\cdot)\in U_{ad},\ z(t_0)=z(t_1)=0\}$$
$$=0.$$

现根据该定理来推导最大值原理.

由定理 II-8-2, 存在 $\lambda_0\leqslant 0$ 和 $\lambda\in\mathbb{R}^n$, 使得

$$\langle\lambda,\ \Phi(t_1,\ t)\{f(t,\ x(t),\ u(t))-f(t,\ x(t),\ u(t))\}\rangle$$

$$+\lambda_0 b^0(t,\ u(t))$$

$$=\max_{v\in U}\{\langle\lambda,\ \Phi(t_1,\ t)\{f(t,\ x(t),\ v)-f(t,\ x(t),\ u(t))\}\rangle$$

$$+\lambda_0 b^0(t,\ v)\},$$

即

$$\lambda_0 f^0(t,\ x(t),\ u(t))$$

$$+\langle-\lambda_0\tilde{\psi}(t)+\Phi^T(t_1,\ t)\lambda,\ f(t,\ x(t),\ u(t))\rangle$$

$$=\max_{v\in U}\{\lambda_0 f^0(t,\ x(t),\ v)$$

$$+\langle-\lambda_0\tilde{\psi}(t)+\Phi^T(t_1,\ t)\lambda,\ f(t,\ x(t),\ v)\rangle\}$$

在$[t_0,\ t_1]$上几乎处处成立.

置 $\psi(\cdot)=-\lambda_0\tilde{\psi}(\cdot)+\Phi^T(t_1,\ \cdot)\lambda,$

那末

$$\frac{d\psi(t)}{dt}=-\lambda_0\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\psi(t),$$

$$\psi(t_1)=\lambda,$$

并且

$$\lambda_0 f^0(t,\ x(t),\ u(t))+\langle\psi(t),\ f(t,\ x(t),\ u(t))\rangle$$

$$=\max_{v\in U}\{\lambda_0 f^0(t,\ x(t),\ v)+\langle\psi(t),\ f(t,\ x(t),v)\rangle\},$$

在$[t_0,\ t_1]$上几乎处处成立.

因此，当值函数$V(\cdot)$具有局部李普希茨条件时，最大值原理成立.

在证明中我们看到，第二章的结果在推导中是重要的.

三 一般情形

引理 1 设$u(\cdot)$和$x(\cdot)=x(\cdot,\ u)$是最优控制问题的解，L是半直线

$$x^0<0,\quad x=0,$$

那末L不含在凸锥W的内部.

证 设

$$e_1,\ e_2,\ \cdots,\ e_n$$

是与 L 正交的互相正交的单位向量.

如果

$$q=(q^0,\ 0)\in L \quad (q^0<0)$$

是 W 的内点,我们将证明 $u(\cdot)$ 不是最优的.

置

$$P=\left\{\sum_{j=1}^{n}\langle x, e_j\rangle e_j \,\middle|\, x\in W, x^0=q^0\right\}.$$

因为 q 是 W 的内点,并且 $q\in L$,所以 $\langle q,\ e_j\rangle=0\ (j=1,\cdots,n)$,从而 O 是以 $e_1,\ e_2,\ \cdots,\ e_n$ 为基的 n 维空间中集 P 的内点.这时,在 P 中存在 $p_1,\ p_2,\cdots,\ p_{n+1}$,使得以 $\{p_j\}$ 为顶点的单纯形 B 含有 O 为内点,即存在

$$x^j\in W \quad (x^{j0}=q^0),$$

使得

$$p_j=\sum_{l=1}^{n}\langle x^j, e_l\rangle e_l \quad (j=1,\ 2,\ \cdots,\ n+1).$$

根据 $x^j\in W$,存在 $u_j^\varepsilon(\cdot)\in U_{ad}$,使得

$$x(t_1,\ u_j^\varepsilon)=x(t_1)+\varepsilon x^j+o(\varepsilon).$$

从而,对于 B 中的点

$$p=\sum_{j=1}^{n+1}\rho_j p_j, \quad \rho_j\geqslant 0, \quad \sum_{j=1}^{n+1}\rho_j=1,$$

存在 $u^\varepsilon(\cdot,\ p)\in U_{ad}$,使得

$$x(t_1, u^\varepsilon(\cdot,\ p))=x(t_1)+\varepsilon\sum_{j=1}^{n+1}\rho_j x^j+o(\varepsilon).$$

设 $T(\cdot,\ \varepsilon):B\to R^n$ 由下式决定:

$$T(p,\ \varepsilon)=\frac{1}{\varepsilon}\sum_{l=1}^{n}\langle x(t_1,\ u^\varepsilon(\cdot,\ p))-x(t_1),\ e_l\rangle e_l.$$

当 $0<\varepsilon<\varepsilon_0$ 时,它是连续的,并且

$$T(p,\ \varepsilon)=\sum_{j=1}^{n+1}\rho_j\sum_{l=1}^{n}\langle x^j, e_l\rangle e_l+o(1)$$

$$= \sum_{j=1}^{n+1} \rho_j \boldsymbol{p}_j + o(1) = \boldsymbol{p} + o(1)$$

因此，不等式

$$\|T(\boldsymbol{p}, \varepsilon) - \boldsymbol{p}\| = o(1) < \|\boldsymbol{p}\|$$

对于 $\boldsymbol{p} \in \partial B$($B$ 的边界)和 $0 < \varepsilon < \varepsilon_0$ 成立.

从而由 $$\tilde{T}(\boldsymbol{p}, \varepsilon) = \boldsymbol{p} - T(\boldsymbol{p}, \varepsilon)$$

决定的 $\tilde{T}(\cdot, \varepsilon): B \to \mathbb{R}^n$，是连续的，并且当 $0 < \varepsilon < \varepsilon_0$ 时，$\tilde{T}(\cdot, \varepsilon): B \to B$. 根据布劳维尔 (L. E. J. Brouwer) 不动点定理 (见附录 3)，存在 $\boldsymbol{p}_0 \in B$，使得

$$\tilde{T}(\boldsymbol{p}_0, \varepsilon) = \boldsymbol{p}_0,$$

即 $$T(\boldsymbol{p}_0, \varepsilon) = \mathbf{0}.$$

从而 $$\langle \boldsymbol{x}(t_1, u^\varepsilon(\cdot, \boldsymbol{p}_0)) - \boldsymbol{x}(t_1), \boldsymbol{e}_l \rangle = 0$$
$$(l = 1, 2, \cdots, n).$$

所以 $$x(t_1, u^\varepsilon(\cdot, \boldsymbol{p}_0)) - x(t_1) = 0,$$

并且

$$x^0(t_1, u^\varepsilon(\cdot, \boldsymbol{p}_0)) = x^0(t_1) + \varepsilon \sum_{j=1}^{n+1} \rho_j^0 q^0 + o(\varepsilon)$$
$$= x^0(t_1) + \varepsilon q^0 + o(\varepsilon),$$

所以 $$x^0(t_1, u^\varepsilon(\cdot, \boldsymbol{p}_0)) < x^0(t_1),$$
$$x(t, u^\varepsilon(\cdot, \boldsymbol{p}_0)) = x(t_1) = x_1. \,]$$

它与 $u(\cdot)$ 的最优性矛盾.

因此，L 上的点不能是 W 的内点. ∎

最大值原理的证明

设 $u(\cdot)$ 和 $x(\cdot) = x(\cdot, u)$ 是最优控制问题的解. 根据引理 1，L 不含在凸锥 W 的内部，$\mathbf{0}$ 是 L 和 W 的公共顶点. 根据凸锥的分离性定理 II-2-8，存在

$$\psi^* = (\psi_0^*, \psi^*) \neq 0,$$

使得超平面

$$\langle \psi^*, x \rangle = 0$$

把 W 和 L 分在两侧，即

$$\langle \psi^*, x \rangle \leqslant 0 \tag{8}$$

对于 $x \in W$ 成立；而

$$\langle \psi^*, x \rangle \geqslant 0$$

对于 $x \in L$ 成立.

因为 $(-1, 0) \in L$，所以由上式得到

$$\psi_0^* \leqslant 0.$$

对于 $v(\cdot) \in U_{ad}$，由方程组

$$\frac{d\delta x^0(t)}{dt} = \left\langle \frac{\partial f^0(t, x(t), u(t))}{\partial x}, \delta x(t) \right\rangle$$
$$+ f^0(t, x(t), v(t)) - f^0(t, x(t), u(t)), \tag{9}$$
$$\frac{d\delta x(t)}{dt} = \left(\frac{\partial f(t, x(t), u(t))}{\partial x} \right)^T \delta x(t)$$
$$+ f(t, x(t), v(t)) - f(t, x(t), u(t)),$$
$$\delta x^0(t_0) = 0,$$
$$\delta x(t_0) = 0$$

决定的 $\delta x(t_1) \in W$，所以由(8)得到

$$\langle \psi^*, \delta x(t_1) \rangle \leqslant 0. \tag{10}$$

设 $\psi(\cdot)$ 是伴随方程

$$\frac{d\psi_0}{dt} = 0,$$

$$\frac{d\psi}{dt} = -\psi_0 \frac{\partial f^0(t, x(t), u(t))}{\partial x} - \frac{\partial f(t, x(t), u(t))}{\partial x} \psi(t)$$

的解，适合条件

$$\psi(t_1) = \psi^*.$$

那末根据第一章 § 5 公式(13)，成立着

$$\langle \psi(t_1), \delta x(t_1) \rangle - \langle \psi(t_0), \delta x(t_0) \rangle$$

$$=\int_{t_0}^{t_1}\langle\psi(t),\ f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t))\rangle dt.$$

注意到 $\delta x(t_0)=0$, 由(10)及上式得到

$$\int_{t_0}^{t_1}\langle\psi(t),\ f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t))\rangle dt\leqslant 0,$$

即

$$\int_{t_0}^{t_1}\mathscr{H}(t,\ \psi(t),\ x(t),\ v(t))dt\leqslant\int_{t_0}^{t_1}\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))dt \tag{11}$$

对于 $v(\cdot)\in U_{ad}$ 成立.

(11)是积分型的最大值原理.

由不等式(11)容易推出,最大值条件

$$\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))=\max_{v\in U}\mathscr{H}(t,\ \psi(t),\ x(t),\ v) \tag{12}$$

在 $[t_0,\ t_1]$ 上几乎处处成立. 请读者自行完成. ▌

习　题

1　利用积分型的最大值条件证明(12).

2　设 W 是 $\mathbb{R}^n$ 中凸锥,L 是半射线, 它们有公共顶点, L 上的点都不是 W 的内点(在 $\mathbb{R}^n$ 的意义下),试证存在超平面把它们分在两侧.

§4　横截条件

设系统的状态方程为

$$\frac{dx}{dt}=f(t,\ x,\ u(t)),\quad u(\cdot)\in U_{ad}, \tag{1}$$

其中 f、U 和 U_{ad} 的意义同前. $t_0,\ t_1\in\mathbb{R}$ 给定.

设 $Q_0\subset\mathbb{R}^n$ 是给定的有界凸闭集, $x_1\in\mathbb{R}^n$ 给定. 对于 $x_0\in Q_0$, $u(\cdot)\in U_{ad}$, 方程组(1)以 $(t_0,\ x_0)$ 为初值的解记为 $x(\cdot)=$

$x(\cdot, x_0, u)$. 即它适合(1)，并且

$$x(t_0, x_0, u)=x_0.$$

如果

$$x(t_1, x_0, u)=x_1,$$

就称 $u(\cdot)$ 把系统(1)从 $x_0(\in Q_0)$迁移到 x_1，或称 $u(\cdot)$ 把系统(1)从 Q_0 迁移到状态 x_1.

再设 $f^0: \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 适合前面同样的条件.

置

$$x^0(t, x_0, u)=\int_{t_0}^{t} f^0(s, x(s, x_0, u), u(s))ds.$$

现讨论始端状态 $x_0\in Q_0$ 的最优控制问题.

最优控制问题 选取 $x_0\in Q_0$ 和 $u(\cdot)\in U_{ad}$，使得

$$x(t_0, x_0, u)=x_0, \quad x(t_1, x_0, u)=x_1,$$

并且不等式

$$x^0(t_1, x_0, u)\leqslant x^0(t_1, \tilde{x}_0, v) \tag{2}$$

对于 $\tilde{x}_0\in Q_0$ 和使得

$$x(t_0, \tilde{x}_0, v)=\tilde{x}_0, \quad x(t_1, \tilde{x}_0, v)=x_1$$

成立的容许控制 $v(\cdot)$ 都成立.

满足上述条件的 x_0、$u(\cdot)$ 和 $x(\cdot)=x(\cdot, x_0, u)$ 称为最优控制问题的解. $u(\cdot)$称为最优控制，$x(\cdot)$称为最优轨线，$x^0(t_1, x_0, u)$ 称为泛函的最优值. 有时也称

$$\boldsymbol{x}(\cdot, x_0, u)=(x^0(\cdot, x_0, u), x(\cdot, x_0, u))$$

为最优轨线. x_0 是最优的初始状态.

显然，如果 $u(\cdot)$ 是把系统(1)从 Q_0 迁移到 x_1 的最优控制，x_0 是最优的初始状态，那末 $u(\cdot)$也是把系统(1)从 x_0 迁移到 x_1 的最优控制问题的解. 因此，$u(\cdot)$ 和 $x(\cdot)$满足最大值原理. 剩下的问题是：最优的初始状态 x_0 应当适合怎样的条件？这就是本节所要讨论的横截条件.

为讨论该问题，我们讨论在初始状态和控制都变动时轨线的变分.

设 $x_0\in Q_0$, $u(\cdot)\in U_{ad}$, $x(\cdot)=x(\cdot, x_0, u)$ 给定.

又设 $\tilde{x}_0\in Q_0$, $v(\cdot)\in U_{ad}$. 置 $\delta x_0=(0, \tilde{x}_0-x_0)$, $u^\varepsilon(\cdot)$ 如 §1 所述，考察

$$x^\varepsilon(\cdot)=x(\cdot, x_0+\varepsilon\delta x_0, u^\varepsilon(\cdot)).$$

可以用 §1 的方法证明：当 $\varepsilon>0$ 充分小时，$x^\varepsilon(\cdot)$ 在 $[t_0, t_1]$ 上存在，并且

$$x^\varepsilon(\cdot)=x(\cdot)+\varepsilon\delta x(\cdot)+o(\varepsilon),$$

其中 $\delta x(\cdot)$ 适合

$$\delta x(t)=\delta x_0+\int_{t_0}^t\left(\frac{\partial f(s, x(s), u(s))}{\partial x}\right)^T\delta x(s)ds$$
$$+\int_{t_0}^t\{f(s, x(s), v(s))-f(s, x(s), u(s))\}ds. \tag{3}$$

由(3)，根据定理 II-6-1，可以证明

定理 1 设 Q_0 是 $\mathbb{R}^n$ 中的凸集，那末

$$K(t)=\{\varepsilon\delta x(t)\,|\,\varepsilon\geqslant 0, \tilde{x}_0\in Q_0, v(\cdot)\in U_{ad}\}$$

是 $\mathbb{R}^{n+1}$ 中的凸锥.

定理 2 设 x_0, $u(\cdot)$ 和 $x(\cdot)$ 是最优控制问题的解，那末半直线

$$L: x^0<0, \ x=0$$

不在 $K(t_1)$ 的内部.

它的证明和 §2 引理 1 类似，读者自己完成.

定理 3 设 x_0, $u(\cdot)$ 和 $x(\cdot)$ 是最优控制问题的解，那末存在 $\psi(\cdot)\neq 0$，适合

$$\frac{d\psi_0}{dt}=0, \quad \psi_0\leqslant 0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(t,\psi(t),x(t),u(t))}{\partial x},$$

使得最大值条件

$$\mathscr{H}(t,\psi(t),x(t),u(t))=\max_{v\in U}\mathscr{H}(t,\psi(t),x(t),v). \tag{4}$$

在$[t_0, t_1]$上几乎处处成立，并且不等式

$$\langle\psi(t_0),\tilde{x}_0-x_0\rangle\leqslant 0 \tag{5}$$

对于$\tilde{x}_0\in Q_0$成立.

不等式(5)称为横截条件.

证 由于L不含在$K(t_1)$的内部，所以存在$\psi^*\neq 0$，使得：当$x\in K(t_1)$时

$$\langle\psi^*, x\rangle\leqslant 0, \tag{6}$$

当$x\in L$时

$$\langle\psi^*, x\rangle\geqslant 0.$$

由上式推知

$$\psi_0^*\leqslant 0.$$

由(6)得到

$$\langle\psi^*, \delta x(t_1)\rangle\leqslant 0. \tag{7}$$

在表达式(3)中，取$v(\cdot)=u(\cdot)$，就得到

$$\delta x(t)=\delta x_0+\int_{t_0}^{t}\left(\frac{\partial f(s,x(s),u(s))}{\partial x}\right)^T\delta x(s)ds,$$

所以
$$\frac{d\delta x(t)}{dt}=\left(\frac{\partial f(t,x(t),u(t))}{\partial x}\right)^T\delta x(t),$$

$$\delta x(t_0)=\delta x_0.$$

设$\psi(\cdot)$适合

$$\frac{d\psi_0}{dt}=0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial f(t,x(t),u(t))}{\partial x}\psi(t),$$

$$\psi(t_1)=\psi^*,$$

根据第一章§5公式(13)，得到

$$\langle\psi(t_1),\ \delta x(t_1)\rangle=\langle\psi(t_0),\ \delta x(t_0)\rangle.$$

由(7)和上式推知

$$\langle\psi(t_0),\ \delta x_0\rangle=\langle\psi(t_0),\ \delta x(t_0)\rangle$$
$$=\langle\psi(t_1),\ \delta x(t_1)\rangle\leqslant 0.$$

它就是横截条件(5).

在表达式(3)中，取 $\delta x_0=0$，$v(\cdot)\in U_{ad}$. 由(7)和§2一样，得到最大值条件(4)在$[t_0,\ t_1]$上几乎处处成立. ∎

再考虑 $x_0\in\mathbb{R}^n$ 给定，$x(t_1)\in Q_1$，$Q_1\subset\mathbb{R}^n$ 给定时的最优控制问题.

当 $Q_1=\mathbb{R}^n$ 时，我们在§2的第一段中已证明了最大值原理和横截条件.

现考虑 Q_1 为 $\mathbb{R}^n$ 中凸集的情形.

如果把 t_1 和 $x_1\in Q_1$ 作为初始时刻和初始状态，t_0 和 x_0 作为终端时刻和终端状态，那末它成为前面讨论过的问题.

设 $u(\cdot)\in U_{ad}$，$x(\cdot,\ u)$ 适合

$$x(t,\ x_1,\ u)=x_1+\int_{t_1}^{t}f(s,\ x(s,\ x_1,\ u),\ u(s))ds$$

记
$$x^0(t,\ x_1,\ u)=\int_{t_1}^{t}f^0(s,\ x(s,\ x_1,\ u),\ u(s))ds,$$

那末
$$x^0(t_0,\ x_1,\ u)=-\int_{t_0}^{t_1}f^0(s,\ x(s,\ x_1,\ u),\ u(s))ds.$$

最优控制问题 选取 $u(\cdot)\in U_{ad}$ 和 $x_1\in Q_1$，使得

$$x(t_0,\ x_1,\ u)=x_0,$$

并且不等式

$$x^0(t_0,\ x_1,\ u)\geqslant x^0(t_0,\ \tilde{x}_1,\ v),$$

对于 $\tilde{x}_1\in Q_1$，$v(\cdot)\in U_{ad}$，$x(t_0,\ \tilde{x}_1,\ v)=x_0$ 成立.

这时，横截条件是

$$\langle\psi(t_1),\ \tilde{x}_1-x_1\rangle\geqslant 0$$

对于 $\tilde{x}_1\in Q_1$ 成立.

一般地,对于把系统(1)从 Q_0 迁移到 Q_1 的最优控制问题,有下述定理.

定理 4 设 $u(\cdot)$是把系统(1)从 Q_0 迁移到 Q_1 的最优控制,$x_0\in Q_0$ 和 $x_1\in Q_1$ 是最优的初始状态和终端状态,$x(\cdot)$ 是最优轨线,那末存在 $\psi(\cdot)\neq 0$,适合

$$\psi_0\leqslant 0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))}{\partial x},\tag{8}$$

使得最大值条件

$$\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))=\max_{v\in U}\mathscr{H}(t,\ \psi(t),\ x(t),\ v)\tag{9}$$

在$[t_0,\ t_1]$上几乎处处成立,并且横截条件

$$\langle\psi(t_0),\ \tilde{x}_0-x_0\rangle\leqslant 0\tag{10}$$

和

$$\langle\psi(t_1),\ \tilde{x}_1-x_1\rangle\geqslant 0\tag{11}$$

对于 $\tilde{x}_0\in Q_0$ 和 $\tilde{x}_1\in Q_1$ 成立.

证 置 $y(\cdot)=x(\cdot)-x(t_1)$,那末

$$\frac{dy(t)}{dt}=f(t,\ y(t)+x(t_1),\ u(t)),$$

$$y(t_0)=x(t_0)-x(t_1),$$

$$y(t_1)=0.$$

因此,如果 $u(\cdot)$、$x(\cdot)$适合(1)及 $x(t_0)\in Q_0$ 和 $x(t_1)\in Q_1$,那末 $y(\cdot)=x(\cdot)-x(t_1)$适合上式,并且 $y(t_0)\in Q_0-Q_1$.

反之,设 $y(\cdot):[t_0,\ t_1]\to\mathbb{R}^n$ 满足下述条件:

$$y(t_1)=0,\ y(t_0)=\tilde{x}_0-\tilde{x}_1,\ \tilde{x}_0\in Q_0,\ \tilde{x}_1\in Q_1,$$

$$\frac{dy(t)}{dt}=f(t,\ y(t)+\tilde{x}_1,\ u(t)), \tag{12}$$

那末 $x(\cdot)=y(\cdot)+\tilde{x}_1$ 是方程组(1)相应于 $u(\cdot)\in U_{ad}$ 的解，并且

$$x(t_0)=\tilde{x}_0\in Q_0,\quad x(t_1)=\tilde{x}_1\in Q_1.$$

现考虑下述带参数 $\tilde{x}_1$ 的最优控制问题：选取 $x_1\in Q_1$ 和 $u(\cdot)\in U_{ad}$，它把系统(12)从状态 Q_0-x_1 迁移到 0，并且

$$J(u,\ x_1)=\int_{t_0}^{t_1}f^0(t,\ y(t)+x_1,\ u(t))dt\leqslant J(v,\ \tilde{x}_1)$$

对于 $\tilde{x}_1\in Q_1$ 和把系统(12)从状态 $Q_0-\tilde{x}_1$ 迁移到 0 的容许控制 $v(\cdot)$ 皆成立。

显然，原来的最优控制问题和上述带参数的最优控制问题是等价的。

置

$$y^0(t,\ v,\ \tilde{x}_1)=\int_{t_0}^{t}f^0(t,\ y(t,\ v,\ \tilde{x}_1)+\tilde{x}_1,\ v(t))dt,$$

$$\boldsymbol{y}=\begin{pmatrix}y^0\\ y\end{pmatrix},\qquad \boldsymbol{y}_0=\begin{pmatrix}0\\ y_0\end{pmatrix}.$$

那末

$$\begin{aligned}&\frac{d\boldsymbol{y}}{dt}=\boldsymbol{f}(t,\ y+\tilde{x}_1,\ v(t))dt\\ &\boldsymbol{y}(t_0)=\boldsymbol{y}_0.\end{aligned} \tag{13}$$

设 $x_1\in Q_1$, $u(\cdot)\in U_{ad}$, $y_0\in Q_0-x_1$ 和 $y(\cdot)=y(\cdot,\ y_0,\ u,\ x_1)$ 是带参数的最优控制问题的解. 对于 $\tilde{x}_1\in Q_1$, $\tilde{y}_0\in Q_0-\tilde{x}_1$, $v(\cdot)\in U_{ad}$, 根据§1的方法构造 $u^\varepsilon(\cdot)$，这里 $\varepsilon>0$. 考虑方程组(13)对应于 $x_1+\varepsilon\delta x_1=x_1+\varepsilon(\tilde{x}_1-x_1)$ 和 $u^\varepsilon(\cdot)$ 的以 $(t_0,\ \boldsymbol{y}_0+\varepsilon\delta\boldsymbol{y}_0)=(t_0,\ \boldsymbol{y}_0+\varepsilon(\tilde{\boldsymbol{y}}_0-\boldsymbol{y}_0))$ 为初值的解 $\boldsymbol{y}^\varepsilon(\cdot)$. 和§1一样，可以证明：当 $\varepsilon>0$ 充分小时 $\boldsymbol{y}^\varepsilon(\cdot)$ 在 $[t_0,\ t_1]$ 上存在，并且

$$\boldsymbol{y}^\varepsilon(\cdot)=\boldsymbol{y}(\cdot)+\varepsilon\delta\boldsymbol{y}(\cdot)+o(\varepsilon),$$

其中 $\delta\boldsymbol{y}(\cdot)$ 适合

$$\delta y(t)=\delta y_0+\int_{t_0}^{t}\left(\frac{\partial f(s,\ y(s)+x_1,\ u(s))}{\partial x}\right)^T(\delta x_1+\delta y(s))ds$$

$$+\int_{t_0}^{t}\{f(s,\ y(s)+x_1,\ v(s))$$

$$-f(s,\ y(s)+x_1,\ u(s))\}ds.$$

由上式容易证明:

$$K(t_1)=\{\varepsilon\delta y(t_1)\,|\,\varepsilon\geqslant 0,\ \tilde{x}_1\in Q_1,\ \tilde{y}_0\in Q_0-\tilde{x}_1,\ v(\cdot)\in U_{ad}\}$$

是 $\mathbb{R}^{n+1}$ 中的凸锥.

并且运用引理1的方法可以证明半直线 L 不含在凸锥 $K(t_1)$ 的内部. 所以存在 $\psi^*\neq 0$, 使得

$$\langle\psi^*,\ \delta y(t_1)\rangle\leqslant 0,$$

$$\langle\psi^*,\ (-1,\ 0)\rangle\geqslant 0.$$

从而

$$\psi_0^*\leqslant 0.$$

如果 $\psi(\cdot)$ 是伴随方程

$$\frac{d\psi_0}{dt}=0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial f(t,\ y(t)+x_1,\ u(t))}{\partial x}\psi(t)$$

的解, 满足条件

$$\psi(t_1)=\psi^*.$$

那末根据第一章§5公式(13), 成立着

$$\langle\psi(t_1),\ \delta y(t_1)\rangle-\langle\psi(t_0),\ \delta y(t_0)\rangle$$

$$=\int_{t_0}^{t_1}\langle\psi(t),\ f(t,\ y(t)+x_1,\ v(t))$$

$$-f(t,\ y(t)+x_1,\ u(t))\rangle dt$$

$$+\int_{t_0}^{t_1}\left\langle\psi(t),\ \left(\frac{\partial f(t,\ y(t)+x_1,\ u(t))}{\partial x}\right)^T\delta x_1\right\rangle dt. \qquad (14)$$

但是

$$\int_{t_0}^{t_1}\left\langle\psi(t),\ \left(\frac{\partial f(t,\ y(t)+x_1,\ u(t))}{\partial x}\right)^T\delta x_1\right\rangle dt$$

$$-\left\langle\int_{t_0}^{t_1}\frac{\partial f(t,\ y(t)+x_1,\ u(t))}{\partial x}\psi(t)dt,\ \delta x_1\right\rangle$$

$$=\langle\psi(t_0)-\psi(t_1),\ \delta x_1\rangle. \tag{15}$$

由(14)和(15)得到

$$\langle\psi^*,\ \delta y(t_1)\rangle-\langle\psi(t_0),\ \delta y_0\rangle+\langle\psi(t_1)-\psi(t_0),\ \delta x_1\rangle$$

$$=\int_{t_0}^{t_1}\langle\psi(t),\ f(t,\ y(t)+x_1,\ v(t))$$

$$-f(t,\ y(t)+x_1,\ u(t))\rangle dt. \tag{16}$$

因此，当 $\delta x_1=0$, $\delta y_0=0$ 时，得到

$$\int_{t_0}^{t_1}\langle\psi(t),\ f(t,\ y(t)+x_1,\ v(t))-f(t,\ y(t)+x_1,\ u(t))\rangle dt\leqslant 0.$$

但 $x(\cdot)=y(\cdot)+x_1$，所以最大值条件成立.

如果 $v(\cdot)=u(\cdot)$，由(16)及 $\langle\psi^*,\ \delta y(t_1)\rangle\leqslant 0$ 推出

$$\langle\psi(t_0),\ \delta y_0\rangle-\langle\psi(t_1)-\psi(t_0),\ \delta x_1\rangle\leqslant 0.$$

但 $\delta y_0=\delta x_0-\delta x_1$，所以

$$\langle\psi(t_0),\ \delta x_0\rangle-\langle\psi(t_1),\ \delta x_1\rangle\leqslant 0.$$

取 $\delta x_1=0$ 时，得到

$$\langle\psi(t_0),\ \tilde{x}_0-x_0\rangle\leqslant 0.$$

取 $\delta x_0=0$ 时，得到

$$\langle\psi(t_1),\ \tilde{x}_1-x_1\rangle\geqslant 0.$$

因此，不等式(10)和(11)成立. ▮

习　　题

1　设系统的状态方程为

$$\frac{dx}{dt}=f(t,\ x,\ w,\ u(t)),$$

其中 $f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}^n$, $w\in W\subset\mathbb{R}$, $u(t)\in U\subset\mathbb{R}^m$. 边值条件为

$$x(t_0)=x_0,\qquad x(t_1)=x_1,$$

$(t_0,\ x_0)$和$(t_1,\ x_1)$给定. 性能指标为

$$J(u(\cdot), w)=\int_{t_0}^{t_1} f^0(t, x(t), w, u(t))dt,$$

其中 $f^0: \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$. 试叙述最优控制问题，给出轨线变分方程，推导最优控制 $u(\cdot)$，最优参数 w_0，最优轨线所适合的必要条件.

2 试把上述最优控制问题变换为不含参数的最优控制问题，并导出上述的一切必要条件.

3 用第1题方法讨论定理4所讨论的问题.

§5 时间不固定的最优控制问题

前面所讨论的最优控制问题中，t_1 是给定的. 因而不能包括时间最优控制问题. 现讨论 t_1 不固定的情形.

先讨论自治系统：

$$\frac{dx}{dt}=f(x, u(t)), \tag{1}$$

$$x(t_0)=x_0, \quad x(t_1)=x_1,$$

$$u(t)\in U,$$

这里 $t_0\in\mathbb{R}$, $x_0\in\mathbb{R}^n$ 和 $x_1\in\mathbb{R}^n$ 是给定的，而 t_1 是不固定的；$f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的，并且 $\frac{\partial f}{\partial x}:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^{n\times n}$ 是连续的，$U\subset\mathbb{R}^m$ 给定，U_{ad} 的意义同前.

又设 $f^0:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 和 $\frac{\partial f^0}{\partial x}:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的. 目标泛函(性能指标)为

$$J(u(\cdot))=\int_{t_0}^{t_1} f^0(x(t), u(t))dt. \tag{2}$$

哈密顿函数为

$$\mathscr{H}(\psi, x, u)=\psi_0 f^0(x, u)+\langle\psi, f(x, u)\rangle,$$

伴随方程组是

$$\frac{d\psi_0}{dt}=0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ u(t))}{\partial x}$$
$$=-\psi_0\frac{\partial f^0(x(t),\ u(t))}{\partial x}-\frac{\partial f(x(t),\ u(t))}{\partial x}\psi(t). \tag{3}$$

定理1 设 $u(\cdot)$, $x(\cdot)$ 满足方程组(1), $\boldsymbol{\psi}(\cdot)$ 适合方程组(3), 并且最大值条件

$$\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ u(t))=\max_{v\in U}\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ v) \tag{4}$$

在 $[t_0,\ t_1]$ 上几乎处处成立, 那末

$$\mathscr{M}(\boldsymbol{\psi}(t),\ x(t))=\sup_{v\in U}\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ v)$$

在 $[t_0,\ t_1]$ 上恒为常数.

证 首先证明: $\mathscr{M}(\boldsymbol{\psi}(t),\ x(t))$ 在 $[t_0,\ t_1]$ 上是下半连续的, 即对于 $\bar{t}\in[t_0,\ t_1]$ 和 $\varepsilon>0$, 存在 $\delta>0$, 使得不等式

$$\mathscr{M}(\boldsymbol{\psi}(t),\ x(t))\geqslant\mathscr{M}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}))-\varepsilon$$

对于 $t\in[t_0,\ t_1]$, $|t-\bar{t}|<\delta$ 成立.

事实上, 根据 $\mathscr{M}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}))$ 是 $\mathscr{H}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}),\ \cdot)$ 在 U 的上确界的定义, 存在 $\bar{u}\in U$, 使得

$$\mathscr{H}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}),\ \bar{u})\geqslant\mathscr{M}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}))-\frac{\varepsilon}{2}.$$

因为 $\boldsymbol{\psi}(\cdot)$ 和 $x(\cdot)$ 是连续的, 所以当 $\bar{u}$ 给定时, 函数

$$\mathscr{H}(\boldsymbol{\psi}(\cdot),\ x(\cdot),\ \bar{u})$$

在 $[t_0,\ t_1]$ 上连续, 从而存在 $\delta>0$, 使得不等式

$$\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ \bar{u})\geqslant\mathscr{H}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}),\ \bar{u})-\frac{\varepsilon}{2}$$

对于 $t\in[t_0,\ t_1]$, $|t-\bar{t}|<\delta$ 成立. 因此, 不等式

$$\mathscr{M}(\boldsymbol{\psi}(t),\ x(t))\geqslant\mathscr{H}(\boldsymbol{\psi}(t),\ x(t),\ \bar{u})$$
$$\geqslant\mathscr{M}(\boldsymbol{\psi}(\bar{t}),\ x(\bar{t}))-\varepsilon$$

对于 $t\in[t_0,\ t_1]$, $|t-\bar{t}|<\delta$ 成立. 即 $\mathscr{M}(\boldsymbol{\psi}(\cdot),\ x(\cdot))$ 是下半连

续的.

其次,因为 $u(\cdot)\in U_{ad}$,所以 $u(\cdot):[t_0, t_1]\to U$ 是有界可积的,从而

$$P=\{u \mid u=u(t),\ t\in[t_0, t_1]\}$$

是有界集,它的闭包 $\overline{P}$ 是有界闭集.

记 $$m(\psi, x)=\max_{v\in P}\mathscr{H}(\psi, x, v),$$

那末 $$m(\psi, x)\leqslant\sup_{v\in U}\mathscr{H}(\psi, x, v)=\mathscr{M}(\psi, x).$$

根据条件(4),等式

$$m(\psi(t), x(t))=\mathscr{M}(\psi(t), x(t))$$

在 $[t_0, t_1]$ 上几乎处处成立.

现在证明: $m(\psi(\cdot), x(\cdot))$ 在 $[t_0, t_1]$ 上是绝对连续的.

因为 $\psi(\cdot)$ 和 $x(\cdot)$ 是绝对连续的,所以存在 $M>0$,使得

$$\|\psi(t)\|\leqslant M,\quad \|x(t)\|\leqslant M.$$

记 $$\Omega=\{\psi \mid \|\psi\|\leqslant M\}\times\{x \mid \|x\|\leqslant M\},$$

它是 $\mathbb{R}^{n+1}\times\mathbb{R}^n$ 中的有界闭集. 因为 $\mathscr{H}(\psi, x, u)$ 关于 ψ, x 是连续可微的,所以存在 $K>0$,使得不等式

$$|\mathscr{H}(\psi, x, u)-\mathscr{H}(\tilde{\psi}, \tilde{x}, u)|\leqslant Kd$$

对于 $u\in P$, $(\psi, x)\in\Omega$, $(\tilde{\psi}, \tilde{x})\in\Omega$ 成立,这里

$$d^2=\|\psi-\tilde{\psi}\|^2+\|x-\tilde{x}\|^2.$$

设 u 和 $\tilde{u}$ 满足

$$m(\psi, x)=\mathscr{H}(\psi, x, u),$$

$$m(\tilde{\psi}, \tilde{x})=\mathscr{H}(\tilde{\psi}, \tilde{x}, \tilde{u}),$$

那末 $$\begin{aligned}-Kd&\leqslant\mathscr{H}(\psi, x, \tilde{u})-\mathscr{H}(\tilde{\psi}, \tilde{x}, \tilde{u})\\&\leqslant\mathscr{H}(\psi, x, u)-\mathscr{H}(\tilde{\psi}, \tilde{x}, \tilde{u})\\&\leqslant\mathscr{H}(\psi, x, u)-\mathscr{H}(\tilde{\psi}, \tilde{x}, u)\leqslant Kd,\end{aligned}$$

即 $$|m(\psi, x)-m(\tilde{\psi}, \tilde{x})|\leqslant Kd.$$

所以 $m(\cdot,\ \cdot)$满足李普希茨条件．据据 $\psi(\cdot)$ 和 $x(\cdot)$ 是绝对连续的，推知：$m(\psi(\cdot),\ x(\cdot))$是绝对连续的．

再证明
$$\frac{d}{dt}m(\psi(t),\ x(t))=0$$
在$[t_0,\ t_1]$上几乎处处成立．

事实上，设在时刻t，等式(4)成立，从而
$$m(\psi(t),\ x(t))=\mathscr{H}(\psi(t),\ x(t),\ u(t)).$$
对于 $\bar{t}\neq t$，成立着
$$m(\psi(\bar{t}),\ x(\bar{t}))\geqslant\mathscr{H}(\psi(\bar{t}),\ x(\bar{t}),\ u(t)).$$
所以
$$m(\psi(\bar{t}),\ x(\bar{t}))-m(\psi(t),\ x(t))$$
$$\geqslant\mathscr{H}(\psi(\bar{t}),\ x(\bar{t}),\ u(t))-\mathscr{H}(\psi(t),\ x(t),\ u(t)). \qquad (5)$$

当 $\bar{t}>t$ 时，用 $\bar{t}-t$ 除上式两端，并令 $\bar{t}\to t$，得到
$$\frac{d}{dt}m(\psi(t),\ x(t))\geqslant\left\langle\frac{\partial\mathscr{H}(\psi(t),\ x(t),\ u(t))}{\partial\psi},\ \frac{d\psi(t)}{dt}\right\rangle$$
$$+\left\langle\frac{\partial\mathscr{H}(\psi(t),\ x(t),\ u(t))}{\partial x},\ \frac{dx(t)}{dt}\right\rangle.$$
注意到(1)和(3)，上式右端为0，所以在 $m(\psi(\cdot),\ x(\cdot))$ 的可微点 t 处，成立着
$$\frac{dm(\psi(t),\ x(t))}{dt}\geqslant 0.$$

当 $\bar{t}<t$ 时，用 $\bar{t}-t$ 除(5)的两端，并令 $\bar{t}\to t$，得到
$$\frac{dm(\psi(t),\ x(t))}{dt}\leqslant 0.$$

所以
$$\frac{dm(\psi(t),\ x(t))}{dt}=0$$
在$[t_0,\ t_1]$上几乎处处成立．

因此，$m(\psi(t),\ x(t))$在$[t_0,\ t_1]$上恒为常数．

最后，因为

$$\mathscr{M}(\psi(t),\ x(t)) = m(\psi(t),\ x(t))$$

在$[t_0,\ t_1]$上几乎处处成立，并且$m(\psi(t),\ x(t))$在$[t_0,\ t_1]$上恒为常数和$\mathscr{M}(\psi(\cdot),\ x(\cdot))$是下半连续的，所以$\mathscr{M}(\psi(t),\ x(t))$恒为常数.

定理 2 设$u(\cdot)$和$x(\cdot)$是自治系统(1)最优控制问题的解，那末存在$\psi(\cdot)\neq 0$，适合

$$\frac{d\psi_0}{dt} = 0, \quad \psi_0 \leqslant 0,$$

$$\frac{d\psi(t)}{dt} = -\frac{\partial \mathscr{H}(\psi(t),\ x(t),\ u(t))}{\partial x}, \tag{6}$$

使得

$$\mathscr{H}(\psi(t),\ x(t),\ u(t)) = \mathscr{M}(\psi(t),\ x(t)) \tag{7}$$

在$[t_0,\ t_1]$上几乎处处成立，并且

$$\mathscr{M}(\psi(t),\ x(t)) \equiv 0 \tag{8}$$

在$[t_0,\ t_1]$上处处成立.

证 我们把时间t_1不固定的最优控制问题化为时间固定的最优控制问题.

为此，引进变量τ如下：$t = t_0 + \tau(t_1 - t_0)$，讨论系统

$$\frac{d\bar{x}^0}{d\tau} = f^0(\bar{x},\ \bar{u}(\tau))w, \quad \bar{x}^0(0) = 0,$$

$$\frac{dt}{d\tau} = w, \quad t(0) = t_0, \quad t(1)\text{ 自由},$$

$$\frac{dw}{d\tau} = 0, \quad w(0)\text{ 自由}, \quad w(1)\text{ 自由},$$

$$\frac{d\bar{x}}{d\tau} = f(\bar{x},\ \bar{u}(\tau))w, \quad \bar{x}(0) = x_0, \quad \bar{x}(1) = x_1.$$

它的伴随方程是

$$\frac{d\bar{\psi}_0}{d\tau} = 0, \quad \bar{\psi}_0 \leqslant 0,$$

$$\frac{dd\bar{\psi}_1}{d\tau} = 0,$$

$$\frac{d\bar{\psi}_2}{d\tau}=-\bar{\psi}_0 f^0(\bar{x}(\tau),\ \bar{u}(\tau))-\bar{\psi}_1(\tau)$$

$$-\langle\bar{\psi}(\tau),\ f(\bar{x}(\tau),\ \bar{u}(\tau))\rangle,$$

$$\frac{d\bar{\psi}}{d\tau}=-\bar{\psi}_0\frac{\partial f^0(\bar{x}(\tau),\ \bar{u}(\tau))}{\partial x}w-\frac{\partial f(\bar{x}(\tau),\ \bar{u}(\tau))}{\partial x}\bar{\psi}(\tau)w.$$

这里

$$\bar{u}(\tau)=u(t_0+\tau w),$$

$$\bar{x}(\tau)=x(t_0+\tau w),$$

而 $\bar{\psi}_1(\cdot)$, $\bar{\psi}_2(\cdot)$ 和 $\psi(\cdot)$满足横截条件

$$\bar{\psi}_1(1)=0,$$

$$\bar{\psi}_2(0)=0,\ \bar{\psi}_2(1)=0.$$

因此,我们有

$$\bar{\psi}_0(\tau)\equiv\psi_0\leqslant 0,$$

$$\bar{\psi}_1(\tau)\equiv 0,$$

并且

$$\psi_2(1)=\psi_2(0)+\int_0^1\{-\psi_0 f^0(\bar{x}(\tau),\ \bar{u}(\tau))$$

$$-\langle\psi(\tau),\ f(\bar{x}(\tau),\ \bar{u}(\tau))\rangle\}d\tau,$$

所以

$$\int_{t_0}^{t_1}\{\psi_0 f^0(x(t),\ u(t))+\langle\psi(t),\ f(x(t),\ u(t))\rangle\}dt=0, \tag{9}$$

这里 $\frac{d\psi_0}{dt}=0$,

$$\frac{d\psi(t)}{dt}=-\psi_0\frac{\partial f^0(x(t),\ u(t))}{\partial x}-\frac{\partial f(x(t),\ u(t))}{\partial x}\psi(t).$$

而等式(8)就是

$$\int_{t_0}^{t_1}\mathscr{H}(\psi(t),\ x(t),\ u(t))dt=0. \tag{10}$$

最大值条件是

$$w\bar{\psi}_0(\tau)f^0(\bar{x}(\tau),\ \bar{u}(\tau))+\langle\bar{\psi}(\tau),\ f(\bar{x}(\tau),\ \bar{u}(\tau))\rangle w$$

$$=\max_{v\in U}\{w\bar{\psi}_0(\tau)f^0(\bar{x}(\tau),\ v)+\langle\bar{\psi}(\tau),\ f(\bar{x}(\tau),\ v)\rangle w\},$$

这里 $w=t_1-t_0>0$. 把 τ 换到原来的时间变量 t, 由上式得到最大值条件

$$\psi_0f^0(x(t),\ u(t))+\langle\psi(t),\ f(x(t),\ u(t))\rangle$$
$$=\max_{v\in U}\{\psi_0f^0(x(t),\ v)+\langle\psi(t),\ f(x(t),\ v)\rangle\}.$$

在 $[t_0,\ t_1]$ 上几乎处处成立. 因此

$$\mathscr{H}(\psi(t),\ x(t),\ u(t))=\mathscr{M}(\psi(t),\ x(t)) \qquad (11)$$

在 $[t_0,\ t_1]$ 上几乎处处成立. 根据定理 1, $\mathscr{M}(\psi(\cdot),\ x(\cdot))$ 恒为常数. 再由等式(10)得到 $\mathscr{M}(\psi(t),\ x(t))=0$ 在 $[t_0,\ t_1]$ 上处处成立.

再来讨论非自治系统

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ u(t)),\quad u(t)\in U,$$

目标泛函为

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,\ x(t),\ u(t))dt,$$

边值条件为

$$x(t_0)\in Q_0,\quad x(t_1)\in Q_1,$$

这里 $t_0\in\mathbb{R}$, $Q_0\subset\mathbb{R}^n$, $Q_1\subset\mathbb{R}^n$ 给定, 而 t_1 是不预先固定的. 设 f, f^0, $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$, $\frac{\partial f^0}{\partial t}$, $\frac{\partial f^0}{\partial x}$ 是连续的.

如果引进 $x^{n+1}=t$, 那末系统可写为

$$\frac{dx(t)}{dt}=f(x^{n+1},\ x(t),\ u(t)),$$
$$\frac{dx^{n+1}}{dt}=1,$$

边值条件为

$$x(t_0)\in Q_0,\quad x^{n+1}(t_0)=t_0,$$
$$x(t_1)\in Q_1,\quad x^{n+1}(t_1)\ \text{自由},$$

目标泛函为

$$J(u(\cdot))=\int_{t_0}^{t_1} f^0(x^{n+1}(t),\ x(t),\ u(t))dt.$$

哈密顿函数为

$$H(\psi,\ \psi_{n+1},\ x,\ x^{n+1},\ u)$$
$$=\psi_0 f^0(x^{n+1},\ x,\ u)+\langle\psi,\ f(x^{n+1},\ x,\ u)\rangle+\psi_{n+1},$$

伴随方程为

$$\frac{d\psi_0}{dt}=0,$$

$$\frac{d\psi(t)}{dt}=-\psi_0\frac{\partial f^0(x^{n+1}(t),\ x(t),\ u(t))}{\partial x}$$
$$-\frac{\partial f(x^{n+1}(t),\ x(t),\ u(t))}{\partial x}\psi(t),$$

$$\frac{d\psi_{n+1}}{dt}=-\psi_0\frac{\partial f^0(x^{n+1}(t),\ x(t),\ u(t))}{\partial x^{n+1}}$$
$$-\frac{\partial f(x^{n+1}(t),\ x(t),\ u(t))}{\partial x^{n+1}}\psi(t). \tag{12}$$

横截条件为

$$\langle\psi(t_0),\ \tilde{x}_0-x_0\rangle\leqslant 0 \qquad \forall\tilde{x}_0\in Q_0,$$
$$\langle\psi(t_1),\ \tilde{x}_1-x_1\rangle\geqslant 0 \qquad \forall\tilde{x}_1\in Q_1,$$
$$\psi_{n+1}(t_1)=0. \tag{13}$$

最大值条件为

$$H(\psi(t),\ \psi_{n+1}(t),\ x(t),\ x^{n+1}(t),\ u(t))$$
$$=\max_{v\in U} H(\psi(t),\ \psi_{n+1}(t),\ x(t),\ x^{n+1}(t),\ v)\equiv 0,$$

即

$$\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))+\psi_{n+1}(t)$$
$$=\max_{v\in U}\{\mathscr{H}(t,\ \psi(t),\ x(t),\ v)+\psi_{n+1}(t)\}\equiv 0.$$

因此，根据(12)、(13)和上式，得到

$$\mathscr{H}(t,\ \psi(t),\ x(t),\ u(t))$$

$$=\max_{v\in U}\mathscr{H}(t,\ \psi(t),\ x(t),\ v)=-\psi_{n+1}(t)$$
$$=\int_{t_1}^{t}\frac{\partial\mathscr{H}(s,\ \psi(s),\ x(s),\ u(s))}{\partial t}ds.$$

所以

$$\max_{v\in U}\mathscr{H}(t_1,\ \psi(t_1),\ x(t_1),\ v)=-\psi_{n+1}(t_1)=0.$$

这是用来确定 t_1 的条件．它是对时间不固定时所得到的必要条件．

在上述推导中，假定了 $\frac{\partial f^0}{\partial t}$，$\frac{\partial f}{\partial t}$ 的连续性．这是因为在上述讨论中引进了 $x^{n+1}=t$．如果用§2的方法来直接讨论非自治系统的最优控制问题，f 和 f^0 关于变量 t 的连续性、可微性的要求可以减弱为勒贝格可积性．请读者自行完成．

习　　题

1　利用§1，§2的方法直接讨论最优控制问题：系统为

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ u(t)),\quad u(t)\in U,$$

边值条件为

$$x(t_0)=x_0,\quad x(t_1)=x_1,$$

其中 t_0，x_0 和 x_1 给定，t_1 不预先给定，性能指标为

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,\ x(t),\ u(t))dt,$$

假设 f 和 f^0 适合引言中所述的条件．

2　试述时间固定的最优控制问题可以化为时间不固定的最优控制问题．

3　叙述并讨论带有参数 w 的最优控制问题：系统的状态方程为

$$\frac{dx}{dt}=f(t,\ x,\ w,\ u(t)),\quad u(t)\in U,\quad w\in W,$$
$$x(t_0)=x_0,\quad x(t_1)=x_1,$$

其中 t_0，x_0，x_1 给定．性能指标为

$$J(u(\cdot),\ w)=\int_{t_0}^{t_1}f^0(t,\ x(t),\ w,\ u(t))dt.$$

给出并证明最优控制 $u(\cdot)$，最优参数 $w_0\in W$ 和最优轨线 $x(\cdot)$ 所适合的必要条件.

§6 时滞系统的最优控制

如果信号在传递过程中只是延迟一段时间而复现，就称它具有时滞效应. 在控制系统中，时滞效应的存在使得控制问题复杂化了.

一 时滞系统

微分差分方程描述的动态系统是时滞系统. 设它的方程为

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ x(t-\theta),\ u(t)),\tag{1}$$

其中 $\theta>0$ 是常数，称为系统的时滞；$f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的，并且

$$\frac{\partial f(t,\ x,\ y,\ u)}{\partial x},\quad \frac{\partial f(t,\ x,\ y,\ u)}{\partial y}$$

也是连续的.

$\mathbb{R}^m$ 中的给定集 U 称为控制区域. 如果 $u(\cdot):[t_0,\ t_1]\to U$ 是逐段连续的，就称 $u(\cdot)$ 是容许控制，记为 $u(\cdot)\in U_{ad}$. 为确定起见，假设 $u(\cdot)$ 在间断点处是左连续的，即

$$u(t)=u(t-0).$$

从表达式(1)看出，为确定 $\dot{x}(t)$，不仅需要 t，$x(t)$ 和 $u(t)$ 的值，而且需要知道 $x(t-\theta)$，即 $x(\cdot)$ 在时刻 $t-\theta$ 的值 $x(t-\theta)$ 对现时的 $\dot{x}(t)$ 还有效用. 因此，为确定(1)的解，需要在 $[t_0-\theta,\ t_0]$ 上给定初值函数 $\varphi(\cdot):[t_0-\theta,\ t_0]\to\mathbb{R}^n$.

设当 $t\in[t_0-\theta,\ t_0]$ 时 $x(t)\equiv\varphi(t)$，那末当 $t\in[t_0,\ t_0+\theta]$ 时 $x(t-\theta)\equiv\varphi(t-\theta)$，因而

$$\frac{dx(t)}{dt}=f(t,\ x(t),\ \varphi(t-\theta),\ u(t)),$$

它是常微分方程，以$(t_0, \varphi(t_0))$为初值可以确定$x(\cdot)$在$[t_0, t_0+\theta]$上的值．如果$x(\cdot)$已在$[t_0, t_0+\theta]$上确定，就可以用上述同样的方法确定$x(\cdot)$在$[t_0+\theta, t_0+2\theta]$上的值．逐步地，可以确定$x(\cdot)$在$[t_0, t_1]$上的值(如果相应的常微分方程的解的存在区间可以的话)．

如果$\varphi(\cdot)$是逐段连续的，$u(\cdot)$是容许控制，那末解$x(\cdot)$是逐段光滑的连续函数．

当 $$f(t, x, y, u)=A(t)x+B(t)y+C(t)u$$

时，(1)成为线性时滞系统

$$\frac{dx(t)}{dt}=A(t)x(t)+B(t)x(t-\theta)+C(t)u(t), \tag{2}$$

假设$A(\cdot):[t_0, t_1]\to\mathbb{R}^{n\times n}$，$B(\cdot):[t_0, t_1]\to\mathbb{R}^{n\times n}$和$C(\cdot):[t_0, t_1]\to\mathbb{R}^{n\times m}$是逐段连续的．

设$\Phi(\cdot, \cdot)$适合

$$\begin{aligned}&\frac{d\Phi(t, s)}{dt}=A(t)\Phi(t, s)+B\Phi(t-\theta, s) \qquad (s\leqslant t),\\&\Phi(s, s)\equiv I,\\&\Phi(t, s)\equiv O \quad (t<s),\end{aligned} \tag{3}$$

那末$\Phi(\cdot, \cdot)$在$t_0\leqslant s\leqslant t\leqslant t_1$上是连续的，并且关于$t$是逐段连续可微的．这时，由下式决定的$x(\cdot)$：

$$\begin{aligned}&x(t)=\int_{t_0}^{t}\Phi(t, s)C(s)u(s)ds \qquad (t_0\leqslant t),\\&x(t)=0 \qquad (t<t_0),\end{aligned} \tag{4}$$

在$[t_0-\theta, t_1]$上是连续的，并且当$t\in(t_0, t_1]$时

$$\frac{dx(t)}{dt}=\int_{t_0}^{t}\frac{d\Phi(t, s)}{dt}C(s)u(s)ds+C(t)u(t).$$

根据(3)得到

$$\frac{dx(t)}{dt}=A(t)x(t)+B(t)x(t-\theta)+C(t)u(t).$$

所以(4)决定的 $x(\cdot)$ 是方程(2)的解. 它是线性微分方程解的常数变易公式在线性微分差分方程组(2)的推广.

方程组

$$\frac{d\psi(t)}{dt}=-A^T(t)\psi(t)-B^T(t+\theta)\psi(t+\theta)\quad(t_0\leqslant t\leqslant t_1),$$

$$\psi(t)=0\quad(t_1<t\leqslant t_1+\theta)\tag{5}$$

称为(2)的共轭方程组.

这时

$$\frac{d}{dt}\langle\psi(t),\ x(t)\rangle$$
$$=\langle\psi(t),\ A(t)x(t)+B(t)x(t-\theta)+C(t)u(t)\rangle$$
$$-\langle A^T(t)\psi(t)+B^T(t+\theta)\psi(t+\theta),\ x(t)\rangle,$$

所以

$$\langle\psi(t_1),\ x(t_1)\rangle-\langle\psi(t_0),\ x(t_0)\rangle$$
$$=\int_{t_0}^{t_1}\langle\psi(t),\ C(t)u(t)\rangle dt+\int_{t_0}^{t_1}\langle\psi(t),\ B(t)x(t-\theta)\rangle dt$$
$$-\int_{t_0}^{t_1}\langle\psi(t+\theta),\ B(t+\theta)x(t)\rangle dt.$$

但是

$$\int_{t_0}^{t_1}\langle\psi(t+\theta),\ B(t+\theta)x(t)\rangle dt=\int_{t_0+\theta}^{t_1}\langle\psi(t),\ B(t)x(t-\theta)\rangle dt,$$

所以成立着

$$\langle\psi(t_1),\ x(t_1)\rangle-\langle\psi(t_0),\ x(t_0)\rangle$$
$$=\int_{t_0}^{t_1}\langle\psi(t),\ C(t)u(t)\rangle dt+\int_{t_0}^{t_0+\theta}\langle\psi(t),\ B(t)x(t-\theta)\rangle dt.\tag{6}$$

如果当 $t\in[t_0-\theta,\ t_0)$ 时 $x(t)=0$, 那末

$$\langle\psi(t_1),\ x(t_1)\rangle-\langle\psi(t_0),\ x(t_0)\rangle$$
$$=\int_{t_0}^{t_1}\langle\psi(t),\ C(t)u(t)\rangle dt.\tag{7}$$

公式(6)和(7)称为(2)和(5)的对偶公式.

二 最优控制问题

设系统的状态方程为(1)，t_0，t_1，$\varphi(\cdot)$和x_1给定.

对于$u(\cdot)\in U_{ad}$，如果方程(1)以$\varphi(\cdot)$为初值函数的解$x(\cdot)$满足

$$x(t_1)=x_1,$$

就称$u(\cdot)$把系统(1)从$\varphi(\cdot)$迁移到x_1. 这时考虑泛函

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,x(t),x(t-\theta),u(t))dt,$$

这里$f^0:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$和$\dfrac{\partial f^0}{\partial x}$，$\dfrac{\partial f^0}{\partial y}$是连续的.

最优控制问题: 选取$u(\cdot)\in U_{ad}$把系统(1)从$\varphi(\cdot)$迁移到x_1，使得

$$J(u(\cdot))\leqslant J(v(\cdot))$$

对于把系统(1)从$\varphi(\cdot)$迁移到x_1的容许控制$v(\cdot)\in U_{ad}$都成立.

置

$$\boldsymbol{x}=\begin{pmatrix}x^0\\x\end{pmatrix},\quad \boldsymbol{f}=\begin{pmatrix}f^0\\f\end{pmatrix},\quad \boldsymbol{\varphi}=\begin{pmatrix}0\\\varphi\end{pmatrix}$$

考虑

$$\begin{gathered}\frac{d\boldsymbol{x}(t)}{dt}=\boldsymbol{f}(t,\ \boldsymbol{x}(t),\ \boldsymbol{x}(t-\theta),\ u(t))\qquad(t_0\leqslant t\leqslant t_1),\\ \boldsymbol{x}(t)=\boldsymbol{\varphi}(t)\qquad(t_0-\theta\leqslant t\leqslant t_0),\\ x(t_1)=x_1.\end{gathered}\tag{8}$$

最优控制问题转化为: 选取$u(\cdot)\in U_{ad}$，使$x^0(t_1,\ u)$达到最小值.

三 轨线的变分

设$\boldsymbol{x}(\cdot)$是(8)相应于$u(\cdot)$的轨线.

又设$v(\cdot)\in U_{ad}$，对于$s>0$，存在$u^s(\cdot)\in U_{ad}$，使得

$$s\int_{t_0}^{t}\{\boldsymbol{f}(s,\ x(s),\ x(s-\theta),\ v(s))$$

$$-\boldsymbol{f}(s, x(s), x(s-\theta), u(s))\}ds$$

$$=\int_{t_0}^{t}\{\boldsymbol{f}(s, x(s), x(s-\theta), u^{\varepsilon}(s))$$

$$-\boldsymbol{f}(s, x(s), x(s-\theta), u(s))\}ds+o(\varepsilon),$$

并且 $$\mu\{t\,|\,u^{\varepsilon}(t)\neq u(t)\}=\varepsilon(t_1-t_0).$$

系统(1)对应于 $u^{\varepsilon}(\cdot)$ 以 $\varphi(\cdot)$ 为初值函数的轨线记为 $\boldsymbol{x}^{\varepsilon}(\cdot)$，那末当 $\varepsilon>0$ 充分小时 $\boldsymbol{x}^{\varepsilon}(\cdot)$ 在 $[t_0, t_1]$ 上存在，并且成立着

$$\boldsymbol{x}^{\varepsilon}(\cdot)=\boldsymbol{x}(\cdot)+\varepsilon\delta\boldsymbol{x}(\cdot)+o(\varepsilon),$$

其中 $\delta\boldsymbol{x}(\cdot)$ 的初值函数为 $\mathbf{0}$，即当 $t\in[t_0-\theta, t_0]$ 时

$$\delta\boldsymbol{x}(t)=\mathbf{0},$$

并且当 $t\in[t_0, t_1]$ 时

$$\frac{d\delta\boldsymbol{x}(t)}{dt}=\left(\frac{\partial\boldsymbol{f}(t, x(t), x(t-\theta), u(t))}{\partial x}\right)^{T}\delta x(t)$$

$$-\left(\frac{\partial\boldsymbol{f}(t, x(t), x(t-\theta), u(t))}{\partial y}\right)^{T}\delta x(t-\theta)$$

$$+\boldsymbol{f}(t, x(t), x(t-\theta), v(t))$$

$$-\boldsymbol{f}(t, x(t), x(t-\theta), u(t)). \tag{9}$$

它的证明与§1类似，读者自行完成.

根据(9)和常数变易公式得到

$$\delta\boldsymbol{x}(t_1)=\int_{t_0}^{t_1}\Phi(t_1, t)\{\boldsymbol{f}(t, x(t), x(t-\theta), v(t))$$

$$-\boldsymbol{f}(t, x(t), x(t-\theta), u(t))\}dt,$$

其中 $\Phi(\cdot, \cdot)$ 适合

$$\frac{d\Phi(t, s)}{dt}=\left(\frac{\partial\boldsymbol{f}(t, x(t), x(t-\theta), u(t))}{\partial \boldsymbol{x}}\right)^{T}\Phi(t, s)$$

$$+\left(\frac{\partial\boldsymbol{f}(t, x(t), x(t-\theta), u(t))}{\partial \boldsymbol{y}}\right)^{T}\Phi(t-\theta, s)$$

$$(s\leqslant t),$$

$\Phi(s, s)=I$

$$\Phi(t, s)=O \qquad (t<s).$$

因此，和§1一样可以证明

定理1 集

$$K(t_1)=\{\varepsilon\delta\boldsymbol{x}(t_1)\mid \varepsilon\geqslant 0,\ v(\cdot)\in U_{ad}\}$$

是 $\mathbb{R}^{n+1}$ 中的凸锥.

同样可证

定理2 设 $u(\cdot)$ 和 $\boldsymbol{x}(\cdot)$ 是最优控制问题的解，那末

$$L:\ x^0<0,\ x=0$$

不含在 $K(t_1)$ 的内部.

四 最大值原理

定理3 设 $u(\cdot)$ 和 $\boldsymbol{x}(\cdot)$ 是最优控制问题的解，那末存在 $\boldsymbol{\psi}(\cdot)\neq\boldsymbol{0}$, $\psi(\cdot)$ 适合

$$\frac{d\psi_0(t)}{dt}=0,\ \psi_0\leqslant 0,$$

$$\frac{d\psi(t)}{dt}=-\frac{\partial f(t, x(t), x(t-\theta), u(t))}{\partial x}\psi(t)$$

$$-\frac{\partial f(t+\theta, x(t+\theta), x(t), u(t+\theta))}{\partial y}\boldsymbol{\psi}(t+\theta),$$

$$(t_0\leqslant t\leqslant t_1), \tag{10}$$

$$\boldsymbol{\psi}(t)=\boldsymbol{0}\quad (t_1<t),$$

使得最大值条件

$$\langle\boldsymbol{\psi}(t), \boldsymbol{f}(t, x(t), x(t-\theta), u(t))\rangle$$

$$=\max_{v\in U}\langle\boldsymbol{\psi}(t), \boldsymbol{f}(t, x(t), x(t-\theta), v)\rangle \tag{11}$$

在 $[t_0, t_1]$ 上成立.

证 根据定理2, L 不含在凸锥 $K(t_1)$ 的内部. 所以存在 $\boldsymbol{\psi}^*\neq\boldsymbol{0}$, 使得

$$\langle\boldsymbol{\psi}^*, \delta\boldsymbol{x}(t_1)\rangle\leqslant 0, \tag{12}$$

$$\left\langle \psi^*, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\rangle \geqslant 0.$$

所以
$$\psi_0^* \leqslant 0.$$

设 $\boldsymbol{\psi}(\cdot)$是(10)的解，适合

$$\boldsymbol{\psi}(t_1) = \boldsymbol{\psi}^*, \ \boldsymbol{\psi}(t) = \mathbf{0} \qquad (t_1 < t).$$

那末

$$\langle \boldsymbol{\psi}(t_1), \ \delta \boldsymbol{x}(t_1) \rangle \leqslant 0. \tag{13}$$

根据对偶关系式，有

$$\langle \boldsymbol{\psi}(t_1), \ \delta \boldsymbol{x}(t_1) \rangle - \langle \boldsymbol{\psi}(t_0), \ \delta \boldsymbol{x}(t_0) \rangle$$
$$= \int_{t_0}^{t_1} \langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ v(t))$$
$$- \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ u(t)) \rangle dt.$$

因为 $\delta x(t) \equiv 0 \ (t \in [t_0 - \theta, \ t_0])$，所以由上式及(13)得到

$$\int_{t_0}^{t_1} \langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ v(t))$$
$$- \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ u(t)) \rangle dt \leqslant 0.$$

由此得证：当 t 是 $u(\cdot)$的连续点时，成立着

$$\langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ v) \rangle$$
$$\leqslant \langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ u(t)) \rangle$$

即最大值条件成立.

对于 $u(\cdot)$的间断点 t，$\bar{t} < t$ 是 $u(\cdot)$ 的连续点，所以

$$\langle \boldsymbol{\psi}(\bar{t}), \boldsymbol{f}(\bar{t}, \ x(\bar{t}), \ x(\bar{t}-\theta), \ v \rangle$$
$$\leqslant \langle \boldsymbol{\psi}(\bar{t}), \boldsymbol{f}(\bar{t}, \ x(\bar{t}), \ x(\bar{t}-\theta), \ u(\bar{t})) \rangle$$

令 $\bar{t} \to t-0$，得到

$$\langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ v \rangle$$
$$\leqslant \langle \boldsymbol{\psi}(t), \boldsymbol{f}(t, \ x(t), \ x(t-\theta), \ u(t-\theta)) \rangle$$

但 $u(t-0) = u(t)$，所以对于 $u(\cdot)$ 的间断点 t，最大值条件也成立. ▌

五　线性时滞系统的二次最优控制

设系统的状态方程是线性微分差分方程 (2)，容许控制为 $u(\cdot):[t_0, t_1]\to\mathbb{R}^m$ 是平方可积的.

再设当 $t\in[t_0, t_1]$ 时

$$Q(t)\geqslant 0, \quad R(t)>0,$$

并且 $Q(\cdot)$, $R(\cdot)$ 是连续的. 考虑目标泛函

$$J(u(\cdot))=\frac{1}{2}\int_{t_0}^{t_1}\{\langle Q(t)x(t), x(t)\rangle+\langle R(t)u(t), u(t)\rangle\}dt$$

取最小值的最优控制问题.

置

$$f^0(t, x, y, u)=\frac{1}{2}\{\langle Q(t)x, x\rangle+\langle R(t)u, u\rangle\},$$

哈密顿函数是

$$\begin{aligned}&\mathscr{H}(t, \psi, x, y, u)\\&\quad=\psi_0/2\cdot\{\langle Q(t)x, x\rangle+\langle R(t)u, u\rangle\}\\&\quad\quad+\langle\psi, A(t)x+B(t)y+C(t)u\rangle,\end{aligned}$$

伴随方程成为

$$\frac{d\psi_0}{dt}=0,$$

$$\frac{d\psi(t)}{dt}=-\psi_0 Q(t)x(t)-A^T(t)\psi(t)-B^T(t+\theta)\psi(t+\theta).$$

由于 $\mathscr{H}$ 在 $u(t)$ 处达到最大值，所以

$$\begin{aligned}&\frac{\partial\mathscr{H}(t, \psi(t), x(t), x(t-\theta), u(t))}{\partial u}\\&\quad=\psi_0 R(t)u(t)+C^T(t)\psi(t)=0.\end{aligned}$$

因为 $x(t_1)$ 是自由的，所以 $\psi(t_1)=0$. 从而 $\psi_0=0$ 是不容许的，取 $\psi_0=-1$. 那末伴随方程成为

$$\psi_0=-1,$$

$$\frac{d\psi(t)}{dt}=Q(t)x(t)-A^T(t)\psi(t)-B^T(t+\theta)\psi(t+\theta),$$

并且 $u(\cdot)$ 适合

$$u(t)=R^{-1}(t)C^{T}(t)\psi(t).$$

进而，可以由

$$\frac{dx(t)}{dt}=A(t)x(t)+B(t)x(t-\theta)+C(t)R^{-1}(t)C^{T}(t)\psi(t),$$

$$\frac{d\psi(t)}{dt}=Q(t)x(t)-A^{T}(t)\psi(t)-B^{T}(t+\theta)\psi(t+\theta),$$

$$x(t)=\varphi(t)\quad(t_0-\theta\leqslant t\leqslant t_0),$$

$$\psi(t)=0\quad(t_1\leqslant t)$$

讨论 $\psi(\cdot)$ 和 $x(\cdot)$ 之间的关系，以建立状态反馈. 这是一件复杂的事，有兴趣的读者可以查阅有关文献，这里不讨论了.

还应指出，时滞系统的状态空间并不是有限维空间 $\mathbb{R}^n$，而是无限维空间 $C([-\theta, 0]; \mathbb{R}^n)$. 这是因为，只有给定了初值函数 $\varphi(\cdot)$，才能确定它的解. 因此，在时滞系统的最优控制问题中，应当讨论终端是 $\psi(\cdot)\in C([-\theta, 0]; \mathbb{R}^n)$ 的情形，而不应当仅讨论 $x(t_1)=x_1\in\mathbb{R}^n$ 的情形. 这是时滞最优控制问题的本质所在. 但是，这时问题将变得十分复杂，最大值原理是否成立还需讨论.

§7 艾克兰变分原理与最大值原理的证明

在最大值原理的证明中，轨线的变分、布劳维尔不动点定理和凸锥的分离性定理起着重要的作用. 由于最大值原理的重要性，人们试图简化它的证明. 利用动态规划方法的最新发展的简化证明，由审校者在附录中给出. 这里讲述法国数学家艾克兰(I. Ekeland)的变分原理，并利用该变分原理给出最大值原理的证明.

为讲述艾克兰变分原理，我们需要距离空间的知识. 它可在夏道行等的书[3]中找到. 这里简述之.

设 V 是一非空的集．假如对于 V 中任意的元素 u、v，给定一个实数 $d(u, v)$ 与之对应，且 $d(\cdot, \cdot)$ 适合

1° $d(u, v)\geqslant 0$，而且 $d(u, v)=0$ 的充要条件是 $u=v$；

2° 三点不等式：

$$d(u, v)\leqslant d(u, w)+d(v, w) \qquad (w\in V). \tag{1}$$

就称 $d(u, v)$ 为 u，v 之间的距离．而称 V 按距离 d 成为距离空间．

由(1)，当 $w=u$ 时，根据 1° 得到

$$d(u, v)\leqslant d(v, u).$$

同样，u，v 互换得到

$$d(v, u)\leqslant d(u, v).$$

从而 d 具有对称性，即

3° $d(u, v)=d(v, u)$．

在距离空间 V 中，如果 u_n、$u\in V$，满足

$$\lim_{n\to\infty} d(u_n, u)=0,$$

就称 u 是 $\{u_n\}$ 的极限点．

显然，如果 $\{u_n\}$ 以 u 为极限，那末

$$\lim_{n, m\to\infty} d(u_n, u_m)=0. \tag{2}$$

但是，由(2)成立不一定能推知 $\{u_n\}$ 有极限．如果距离空间 V 中的任一适合(2)的点列 $\{u_n\}$ 必有极限，则称 V 是一个完备的距离空间．

设 $F:V\to(-\infty, +\infty)$，如果对于 $v\in V$，$\varepsilon>0$，存在 $\delta>0$，使得当 $d(u, v)<\delta$ 时

$$F(u)>F(v)-\varepsilon,$$

就称 F 在 v 处是下半连续的．当 F 在 V 的每一点是下半连续的，就称 F 是下半连续的．

艾克兰变分原理 设 V 是一个完备的距离空间，F 是 V 上的下半连续函数，且 F 在 V 上是下方有界的，即

$$\inf_{v\in V} F(v)$$

有限，设 $\varepsilon>0$ 和 $u\in V$ 适合

$$F(u)\leqslant\inf\{F(v)\,|\,v\in V\}+\varepsilon, \tag{3}$$

那末存在 $v\in V$，使得

i) $F(v)\leqslant F(u)$;

ii) $d(v,\ u)\leqslant\lambda$;

iii) 对于 $w\in V$，成立着

$$F(w)\geqslant F(v)-\frac{\varepsilon}{\lambda}d(w,\ v), \tag{4}$$

这里 $\lambda>0$.

证 取 $u_0=u$，如果(4)对于 $v=u_0$ 成立，我们证毕。否则，存在 $w\in V$，使得

$$F(w)<F(u_0)-\frac{\varepsilon}{\lambda}d(w,\ u_0).$$

记
$$S_0=\left\{w\in V\,|\,w\neq u_0,\ F(w)<F(u_0)-\frac{\varepsilon}{\lambda}d(w,\ u_0)\right\},$$

它是非空的。由 S_0 的定义知：当 $w\in S_0$ 时

$$0<\frac{\varepsilon}{\lambda}d(w,\ u_0)<F(u_0)-F(w)$$
$$\leqslant F(u_0)-\inf\{F(w)\,|\,w\in S_0\}.$$

又根据下确界的定义，存在 $u_1\in S_0$，使得

$$0\leqslant F(u_1)-\inf\{F(w)\,|\,w\in S_0\}$$
$$\leqslant\frac{1}{2}(F(u_0)-\inf\{F(w)\,|\,w\in S_0\}).$$

这时
$$\frac{\varepsilon}{\lambda}d(u_1,\ u_0)<F(u_0)-F(u_1).$$

如果当 $v=u_1$ 时，(4)对 $w\in S_0$ 成立，而当 $w\bar{\in} S_0$ 时，成立着

$$F(w) \geqslant F(u_0) - \frac{\varepsilon}{\lambda} d(w, u_0),$$

$$\geqslant F(u_0) - \frac{\varepsilon}{\lambda} d(u_1, u_0) - \frac{\varepsilon}{\lambda} d(w, u_1)$$

$$\geqslant F(u_1) - \frac{\varepsilon}{\lambda} d(w, u_1).$$

从而，(4)对 $w \in V$ 皆成立．证毕．否则，集

$$S_1 = \left\{ w \in S_0 \,|\, w \neq u_1,\ F(w) < F(u_1) - \frac{\varepsilon}{\lambda} d(w, u_1) \right\}$$

非空．运用前面同样的方法，存在

$$u_2 \in S_1,$$

使得

$$0 \leqslant F(u_2) - \inf\{F(w) \,|\, w \in S_1\}$$

$$\leqslant \frac{1}{2} (F(u_1) - \inf\{F(w) \,|\, w \in S_1\}),$$

$$\frac{\varepsilon}{\lambda} d(u_2, u_1) < F(u_1) - F(u_2).$$

从而，我们可以逐次定义 S_n, u_n，它们满足：

$$S_n = \left\{ w \in S_{n-1} \,|\, w \neq u_n,\ F(w) < F(u_n) - \frac{\varepsilon}{\lambda} d(w, u_n) \right\},$$

$$0 < \frac{\varepsilon}{\lambda} d(u_{n+1}, u_n) < F(u_n) - F(u_{n+1}),\ u_{n+1} \in S_n, \tag{5}$$

$$0 \leqslant F(u_{n+1}) - \inf\{F(w) \,|\, w \in S_n\}$$

$$\leqslant \frac{1}{2} (F(u_n) - \inf\{F(w) \,|\, w \in S_n\}). \tag{6}$$

从而 $\{F(u_n)\}$ 是单调降的序列，根据假设它是下方有界的，所以

$$\lim_{n \to \infty} F(u_n)$$

存在．而由(5)得到：当 $m > n$ 时

$$0 < \frac{\varepsilon}{\lambda} d(u_m, u_n) < F(u_n) - F(u_m), \tag{7}$$

从而 $\lim\limits_{n, m \to +\infty} d(u_m, u_n) = 0$，由 V 的完备性，存在 $v \in V$，使得

$$\lim_{n\to\infty} d(v,\ u_n)=0.$$

又根据 F 的下半连续性，得到

$$F(v)\leqslant\varliminf_{n\to\infty} F(u_n)\leqslant F(u_0)=F(u).$$

在(7)中令 $m\to+\infty$，得到

$$0\leqslant\frac{\varepsilon}{\lambda}d(v,\ u_n)\leqslant F(u_n)-F(v). \tag{8}$$

又由(7)得到

$$0<\frac{\varepsilon}{\lambda}d(u_m,\ u_0)<F(u_0)-F(v_m),$$

即

$$\begin{aligned}0<\frac{\varepsilon}{\lambda}d(u_m,\ u)&<F(u)-F(v_m)\\ &\leqslant F(u)-\inf\{F(w)\,|\,w\in V\}\leqslant\varepsilon,\end{aligned}$$

在上式令 $m\to+\infty$，得到

$$\frac{\varepsilon}{\lambda}d(v,\ u)\leqslant\varepsilon,$$

从而

$$d(v,\ u)\leqslant\lambda.$$

剩下来证明：不等式(4)成立．否则，存在 $\tilde{w}\neq v$，使得

$$F(\tilde{w})<F(v)-\frac{\varepsilon}{\lambda}d(\tilde{w},\ v), \tag{9}$$

把它与不等式(8)相结合，得到

$$F(\tilde{w})<F(u_n)-\frac{\varepsilon}{\lambda}d(\tilde{w},\ u_n),\quad n=0,\ 1,\ 2,\ \cdots.$$

所以 $\tilde{w}\neq u_n$，$n=0,\ 1,\ 2,\ \cdots$．由上式逐次得到

$$\tilde{w}\in S_n,\quad n=0,\ 1,\ 2,\ \cdots.$$

而由不等式(6)得到

$$\begin{aligned}2F(u_{n+1})&\leqslant F(u_n)+\inf\{F(w)\,|\,w\in S_n\}\\ &\leqslant F(u_n)+F(\tilde{w}).\end{aligned}$$

由上式令 $n\to+\infty$ 时得到

$$2\varlimsup_{n\to+\infty} F(u_{n+1}) \leqslant \varlimsup_{n\to+\infty} F(u_n) + F(\widetilde{w}),$$

从而

$$F(v) \leqslant \varliminf_{n\to\infty} F(u_n) \leqslant F(\widetilde{w}) < F(v) - \frac{\varepsilon}{\lambda} d(\widetilde{w}, v).$$

这是不可能的. 因此, 不等式(4)成立. ∎

注 不等式(4)表明: 函数

$$F(w) + \frac{\varepsilon}{\lambda} d(w, v)$$

在 $w=v$ 处达到最小值 $F(v)$.

现讨论 §1 的最优控制问题. 为简单计, 不妨设 $t_0=0$, $t_1=1$. U_{ad} 是取值为 U 的有界可测函数全体, 在其中引进距离

$$d(u, v) = \mu\{t \mid t\in[0, 1], u(t)\neq v(t)\},$$

那末在此距离下 U_{ad} 是一距离空间. 当 U 是有界闭集时, U_{ad} 是完备的距离空间(读者自行证明之).

设 $u(\cdot)$, $x(\cdot)$ 是最优控制问题的解, $v(\cdot)\in U_{ad}$, $\varepsilon>0$, 引进

$$F_\varepsilon(v(\cdot)) = \{\|x(1, v) - x_1\|^2$$

$$+ (\max(0, x^0(1, v) - x^0(1) + \varepsilon))^2\}^{\frac{1}{2}}.$$

容易证明: $F_\varepsilon: U_{ad}\to\mathbb{R}$ 是连续的, 并且

$$F_\varepsilon(v(\cdot)) \geqslant 0, \quad F_\varepsilon(u(\cdot)) = \varepsilon,$$

即

$$F_\varepsilon(u(\cdot)) \leqslant \inf F_\varepsilon(v(\cdot)) + \varepsilon.$$

根据艾克兰变分原理, 存在 $u^\varepsilon(\cdot)\in U_{ad}$, 使得

$$F_\varepsilon(u^\varepsilon(\cdot)) \leqslant F_\varepsilon(u(\cdot)),$$

$$d(u^\varepsilon(\cdot), u(\cdot)) \leqslant \varepsilon^{\frac{1}{2}},$$

并且对于 $w(\cdot)\in U_{ad}$, 成立着

$$F_\varepsilon(w(\cdot)) \geqslant F_\varepsilon(u^\varepsilon(\cdot)) - \sqrt{\varepsilon}\, d(w(\cdot), u^\varepsilon(\cdot)).$$

记 $x^\varepsilon(\cdot) = x(\cdot, u^\varepsilon)$, 对于 $v(\cdot)\in U_{ad}$, $\rho\in(0, 1)$, 取 $E_\rho^\varepsilon\subset[0,1]$, 使得 $\mu(E_\rho^\varepsilon)=\rho$, 并且

$$\rho\int_0^t\begin{pmatrix}f^0(s,\ x^\varepsilon(s),\ v(s))-f^0(s,\ x^\varepsilon(s),\ u^\varepsilon(s))\\ f(s,\ x^\varepsilon(s),\ v(s))-f(s,\ x^\varepsilon(s),\ u^\varepsilon(s))\end{pmatrix}ds$$
$$=\int_{[0,t]\cap E_\rho^\varepsilon}\begin{pmatrix}f^0(s,\ x^\varepsilon(s),\ v(s))-f^0(s,\ x^\varepsilon(s),\ u^\varepsilon(s))\\ f(s,\ x^\varepsilon(s),\ v(s))-f(s,\ x^\varepsilon(s),\ u^\varepsilon(s))\end{pmatrix}ds$$
$$+o(\rho).$$

置
$$u_\rho^\varepsilon(t)=\begin{cases}v(t), & t\in E_\rho^\varepsilon,\\ u^\varepsilon(t), & t\in[0,\ 1]\setminus E_\rho^\varepsilon,\end{cases}$$
那末 $u_\rho^\varepsilon(\cdot)\in U_{ad}$，并且
$$F_\varepsilon(u_\rho^\varepsilon(\cdot))\geqslant F_\varepsilon(u^\varepsilon(\cdot))-\sqrt{\varepsilon}\,d(u_\rho^\varepsilon(\cdot),\ u^\varepsilon(\cdot))$$
$$\geqslant F_\varepsilon(u^\varepsilon(\cdot))-\sqrt{\varepsilon}\,\rho. \tag{10}$$

现分几种情形进行讨论.

1　存在 $\rho_0>0$，使得当 $\rho\in(0,\ \rho_0)$ 时 $x^0(1,\ u_\rho^\varepsilon)-x^0(1)+\varepsilon\geqslant 0$，这时
$$x^0(1,\ u_\rho^\varepsilon)-x^0(1)+\varepsilon\geqslant 0;$$
$$\|x(1,\ u_\rho^\varepsilon)-x_1\|^2-\|x(1,\ u^\varepsilon)-x_1\|^2$$
$$=\langle x(1,\ u_\rho^\varepsilon)+x(1,\ u^\varepsilon)-2x_1,\ x(1,\ u_\rho^\varepsilon)-x(1,\ u^\varepsilon)\rangle$$
$$=2\langle x(1,\ u^\varepsilon)-x_1,\ x(1,\ u_\rho^\varepsilon)-x(1,\ u^\varepsilon)\rangle+o(\rho).$$
如果 $\delta x(\cdot,\ u^\varepsilon)$ 适合
$$\frac{d}{dt}\delta x(t,\ u^\varepsilon)$$
$$=\left(\frac{\partial f(t,\ x^\varepsilon(t),\ u^\varepsilon(t))}{\partial x}\right)^T\delta x(t,\ u^\varepsilon)$$
$$+f(t,\ x^\varepsilon(t),\ v(t))-f(t,\ x^\varepsilon(t),\ u(t)),$$
$$\delta x(0,\ u^\varepsilon)=0, \tag{11}$$
那末
$$x(1,\ u_\rho^\varepsilon)=x(1,\ u^\varepsilon)+\rho\delta x(1,\ u^\varepsilon)+o(\rho),$$
从而
$$\|x(1,\ u_\rho^\varepsilon)-x_1\|^2-\|x(1,\ u^\varepsilon)-x_1\|^2$$
$$=2\langle x(1,\ u^\varepsilon)-x_1,\ \delta x(1,\ u^\varepsilon)\rangle\rho+o(\rho).$$

而

$$(x^0(1,\ u_\rho^\varepsilon)-x^0(1,\ u)+\varepsilon)^2-(x^0(1,\ u^\varepsilon)-x^0(1,\ u)+\varepsilon)^2$$
$$=2\rho(x^0(1,\ u^\varepsilon)-x^0(1,\ u)+\varepsilon)\cdot\delta x^0(1,\ u^\varepsilon)+o(\rho).$$

所以
$$F_\varepsilon^2(u_\rho^\varepsilon(\cdot))-F_\varepsilon^2(u^\varepsilon(\cdot))$$
$$=2\rho\langle x(1,\ u^\varepsilon)-x_1,\ \delta x(1,\ u^\varepsilon)\rangle$$
$$+2\rho(x^0(1,\ u_\rho^\varepsilon)-x^0(1)+\varepsilon)\delta x^0(1,\ u_\rho^\varepsilon)+o(\rho),$$

从而
$$\lim_{\rho\to 0}\frac{F_\varepsilon^2(u_\rho^\varepsilon(\cdot))-F_\varepsilon^2(u^\varepsilon(\cdot))}{2\rho}$$
$$=\langle x(1,\ u^\varepsilon)-x_1,\ \delta x(1,\ u^\varepsilon)\rangle$$
$$+(x^0(1,\ u^\varepsilon)-x^0(1)+\varepsilon)\delta x^0(1,\ u^\varepsilon),$$

因为 $\lim\limits_{\rho\to 0}F_\varepsilon(u_\rho^\varepsilon(\cdot))=F_\varepsilon(u^\varepsilon(\cdot))$，所以由上式得到

$$\lim_{\rho\to 0}\frac{F_\varepsilon(u_\rho^\varepsilon(\cdot))-F_\varepsilon(u^\varepsilon(\cdot))}{\rho}$$
$$=\frac{1}{F_\varepsilon(u^\varepsilon(\cdot))}\{\langle x(1,\ u^\varepsilon)-x_1,\ \ \delta x(1,\ u^\varepsilon)\rangle$$
$$+(x^0(1,\ u^\varepsilon)-x^0(1)+\varepsilon)\delta x^0(1,\ u^\varepsilon)\}.\tag{12}$$

置
$$h_\varepsilon=\frac{x(1,\ u^\varepsilon)-x_1}{F_\varepsilon(u^\varepsilon(\cdot))},$$
$$h_\varepsilon^0=\frac{x^0(1,\ u^\varepsilon)-x^0(1)+\varepsilon}{F_\varepsilon(u^\varepsilon(\cdot))}>0$$

那末
$$\|h_\varepsilon\|^2+(h_\varepsilon^0)^2=1.$$

(10)和(12)相结合得到

$$\langle h_\varepsilon,\ \delta x(1,\ u^\varepsilon)\rangle+h_\varepsilon^0\delta x^0(1,\ u^\varepsilon)\geqslant-\sqrt{\varepsilon}.\tag{13}$$

由于$(h_\varepsilon^0,\ h_\varepsilon)$落在 $\mathbb{R}^{n+1}$ 中的单位球面上，所以存在收敛子列，它的极限为 $(h^0,\ h)$，$(h^0)^2+\|h\|^2=1$，$h^0\geqslant 0$.

因为 $d(u^\varepsilon,\ u)\leqslant\sqrt{\varepsilon}$，所以当 $\varepsilon\to 0$ 时 $x^\varepsilon(\cdot)\to x(\cdot)$，并且注意到(11)得：$\delta x(\cdot,\ u^\varepsilon)$ 收敛于 $\delta x(\cdot)$，它适合

$$\frac{d}{dt}\delta x(t)=\left(\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\right)^T\delta x(t)$$

$$+f(t, x(t), v(t))-f(t, x(t), u(t)),$$

$$\delta x(0)=0,$$

而 $\delta x^0(t, u^\varepsilon)$ 收敛于 $\delta x^0(t)$，其中

$$\delta x^0(1)=\int_0^1\left(\frac{\partial f^0(t, x(t), u(t))}{\partial x}\right)^T\delta x(t)dt$$

$$+\int_0^1\{f^0(t, x(t), v(t))-f^0(t, x(t), u(t))\}dt.$$

在(13)中令 $\varepsilon\to 0$，得到

$$\langle h, \delta x(1)\rangle+h^0\delta x^0(1)\geqslant 0. \tag{14}$$

设 $\psi_0=-h^0\leqslant 0.$

$$-\frac{d\psi(t)}{dt}=\left(\frac{\partial f(t, x(t), u(t))}{\partial x}\right)\psi(t)-\frac{\partial f^0(t, x(t), u(t))}{\partial x}h^0,$$

$$\psi(1)=-h,$$

那末由(14)和第一章§5公式(13)，得到

$$\int_0^1\{\psi_0 f^0(t, x(t), u(t))+\langle\psi(t), f(t, x(t), u(t)\rangle\}dt$$

$$\geqslant\int_0^1\{\psi_0 f^0(t, x(t), v(t))+\langle\psi(t), f(t, x(t), v(t))\rangle\}dt.$$

由此，得证：最大值条件

$$\psi_0 f^0(t, x(t), u(t))+\langle\psi(t), f(t, x(t), u(t))\rangle$$

$$=\max_{v\subset U}\{\psi_0 f^0(t, x(t), v)+\langle\psi(t), f(t, x(t), v)\rangle\}$$

在[0, 1]上几乎处处成立。

2 存在 $\rho\to 0$ 的子列，使得

$$x^0(1, u_\rho^\varepsilon)-x^0(1)+\varepsilon<0.$$

这时 $x^0(1, u^\varepsilon)-x^0(1)+\varepsilon\leqslant 0$，由 $u(\cdot)$、$x(\cdot)$ 的最优性推知

$$x(1, u^\varepsilon)\neq x_1, \quad x(1, u_\rho^\varepsilon)\neq x_1.$$

从而
$$F_\varepsilon^2(u_\rho^\varepsilon(\cdot))=\|x(1, u_\rho^\varepsilon)-x_1\|^2,$$

$$F_\varepsilon^2(u^\varepsilon(\cdot))=\|x(1, u^\varepsilon)-x_1\|^2,$$

所以

$$F_\varepsilon^2(u_\rho^\varepsilon(\cdot))-F_\varepsilon^2(u^\varepsilon(\cdot))$$
$$=2\rho\langle x(1,\ u^\varepsilon)-x_1,\ \delta x(1,\ u^\varepsilon)\rangle+o(\rho).$$

因此

$$\frac{1}{F_\varepsilon(u^\varepsilon(\cdot))}\langle x(1,\ u^\varepsilon)-x_1,\ \delta x(1,\ u^\varepsilon)\rangle\geqslant-\sqrt{\varepsilon}.$$

取 $h_\varepsilon^0=0$, $h_\varepsilon=(x(1,\ u^\varepsilon)-x_1)/F_\varepsilon(u^\varepsilon(\cdot))$, 那末 $\|h_\varepsilon\|^2=1$. 从而(13)成立. 然后和情形 1 一样推得最大值条件. ▌

习 题

1 设 Q 是 $\mathbb{R}^n$ 中的凸闭集. 试证: 对于 $x\in\mathbb{R}^n$, 存在唯一的点 Px, 满足

$$\|x-Px\|=\inf\{\|x-y\|\,|\,y\in Q\},\ Px\in Q,$$

并且算子 $P:\mathbb{R}^n\to Q$ 适合不等式

$$\|Px-P\tilde{x}\|\leqslant\|x-\tilde{x}\|;$$

如果 $x\in Q$, 那末极限

$$\lim_{\rho\to0}\frac{1}{\rho}(\|x+\rho e-P(x+\rho e)\|^2-\|x-Px\|^2)=2\langle x-Px,\ e\rangle$$

对于 $e\in\mathbb{R}^n$ 皆成立.

2 设 $d(u^\varepsilon,\ u)\to0$ (当 $\varepsilon\to0$ 时), 试证: (11)的解 $\delta x(\cdot,\ u^\varepsilon)$当 $\varepsilon\to0$ 时趋于 $\delta x(\cdot)$, 其中 $\delta x(\cdot)$适合

$$\frac{d}{dt}\delta x(t)=\left(\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\right)^T\delta x(t)$$
$$+f(t,\ x(t),\ v(t))-f(t,\ x(t),\ u(t)),$$
$$\delta x(0)=0.$$

3 试利用艾克兰变分原理证明 $x(t_1)\in Q$ 时最优控制问题的最大值原理和横截条件.

4 证明 U_{ad} 在赋于距离

$$d(u,\ v)=\mu\{t\,|\,t\in[0,\ 1],\ u(t)\neq v(t)\}$$

后是距离空间; 当 U 是闭集时, U_{ad} 是完备的距离空间.

第五章 最优控制的近似计算方法

§1 引 言

由于各工程技术领域、经济管理和资源分配等实际应用部门的需要，随着最优化计算方法和电子计算机的迅速发展，近二十年来，最优控制的计算方法也得到了较大的发展，成为最优控制理论中的一个重要组成部分和解决实际应用问题的一个有力工具.

最优控制计算方法的主要内容是研究求解最优控制问题的各种数值计算方法，并研究其收敛性和收敛速度. 这些内容多数是把变分方法和求解非线性规划最优化问题的一些数值方法和技术加以改造、移植和拓展而得到的. 早在 1960 年，Horn 和 Kelley 就发表了后来在庞特里雅金最大值原理中利用的伴随方程组的梯度法，Breakwell 和 Bryson 等以不同的方式采用 Newton 方法研究了求解最优控制问题，Russell 证明了当系统关于控制为线性时可用罚函数方法求解约束最优控制问题，Balakrishnan 提出了"Balakrishnan ε-方法"，而且推出了庞特里雅金最大值原理，Polak 在最优控制问题的梯度法、可行方向法和梯度投影法方面作出了杰出的工作. 在我国，宫锡芳对最优控制问题的计算方法作过系统的研究；陈祖浩研究了约束最优控制问题的罚函数方法，用统一的理论提供了若干充分和充要条件来处理带罚函数的最优控制问题趋于原最优控制问题，还解决了 Russell 提出的困难问题. 由于连续系统最优控制问题是在无限维空间内讨论的，故最优控制的计算方法远较有限维非线性规划问题的计算方法为复杂，迄今为止，最优控制问题的计算方法的研究，在深度和广度方

面都远不如非线性规划的最优化计算方法的研究，一些最优控制问题的计算方法，如最优控制罚函数方法等，其收敛速度等问题都还未解决.

本章讲述由常微分方程来描述的最优控制问题的一些计算方法的基本理论. 计算方法有两类：一关于无约束最优控制问题；二关于约束最优控制问题. 我们将介绍处理无约束最优控制问题的梯度法和把最优控制问题作为边值问题处理的 Newton 法，并介绍处理约束最优控制问题的罚函数方法. 进一步的研究，读者可参见有关文献.

§2 解无约束最优控制问题的梯度法

设受控系统由微分方程组

$$\left.\begin{aligned}&\frac{dx(t)}{dt}=f(t, x(t), u(t)),\ t\in[t_0, t_1],\\&x(t_0)=x_0.\end{aligned}\right\}\tag{1}$$

来描述，性能指标为

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t, x(t), u(t))dt,\tag{2}$$

此处 f^0: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, f: $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, f^0 和 f 关于 x 和 u 连续可微，关于 t 连续，且对每一个 $(x, u)\in\mathbb{R}^n\times\mathbb{R}^m$, $\frac{\partial f^0}{\partial x}$, $\frac{\partial f^0}{\partial u}$, $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial u}$ 关于 t 连续；又 t_0 和 t_1 已给定.

取 Banach 空间 $L^\infty([t_0, t_1]; \mathbb{R}^m)$ 内的任一元素为容许控制，$L^\infty([t_0, t_1]; \mathbb{R}^m)$ 是 $[t_0, t_1]\to\mathbb{R}^m$ 的可测函数的等价类空间，我们把在 $[t_0, t_1]$ 上几乎处处相等的函数看作是同一个函数. 对 $u(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)$，其范数由下式定义：

$$\|u(\cdot)\|_\infty=\operatorname*{ess\,sup}_{t\in[t_0,t_1]}\|u(t)\|$$

$$\triangleq \inf\{\alpha \mid \text{meas}[t\in[t_0, t_1] \mid \|u(t)\|>\alpha]=0\},$$

可证明(参见关肇直《泛函分析讲义》[4])

$$\|u(\cdot)\|_\infty=\lim_{p\to\infty}\left[\int_{t_0}^{t_1}\|u(t)\|^p dt\right]^{1/p},$$

且空间 $L^\infty([t_0, t_1]; \mathbb{R}^m)$ 是一个线性赋范空间. 我们还定义 $u_i(\cdot)\in L^\infty[t_0, t_1]; \mathbb{R}^m)$, $i=1, 2$, 的数量积为

$$\langle u_1(\cdot), u_2(\cdot)\rangle_\infty=\int_{t_0}^{t_1}\langle u_1(t), u_2(t)\rangle dt,$$

这与 Hilbert 空间 $L^2([t_0, t_1)]; \mathbb{R}^m)$内的数量积相同, 即 $\langle u_1, u_2\rangle_\infty=\langle u_1, u_2\rangle_2$(但 $L^\infty([t_0, t_1); \mathbb{R}^m)$ 并不是在此内积下的 Hilbert 空间!).

最优控制问题 选取容许控制 $u(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)$, 使相应的系统(1)的在 t_0 时刻从点 x_0 出发的轨线 $x(\cdot, u)$致使

$$J(u(\cdot))=\inf\{J(v(\cdot)) \mid x(t_0, v)=x_0,$$
$$v(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)\}$$

这里 t_0, t_1 是给定的, 轨线的左端点固定, 右端点自由.

我们要把最优问题(1), (2)转化为一个无限维空间内的非线性规划问题. 为此, 记

$$J(u(\cdot))=\int_{t_0}^{t_1} f^0(t, x(t, u), u(t))dt, \tag{3}$$

则函数 $J: L^\infty([t_0, t_1]; \mathbb{R}^m)\to\mathbb{R}$. 于是, 最优控制问题(1), (2)就等价于下面的最优化问题了:

$$\text{(UP)}\ \inf\{J(u(\cdot)) \mid u(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)\}. \tag{4}$$

下面, 我们先就求解无约束最优化问题(UP)的算法作一般性的讨论.

在通常的情形下, 求解这个对 $u\in L^\infty\triangleq L^\infty([t_0, t_1]; \mathbb{R}^m)$无约束的最优化问题(UP)的算法, 大都用的是迭代法, 其基本思想是: 从 Banach 空间 L^∞ 内的一个闭子集 C 中某一点 u_0 出发, 找出一

点 u_1 使 $J(u_1)<J(u_0)$，然后再找一个点 u_2，使 $J(u_2)<J(u_1)$，…，即使每进行一步，函数值都有所下降，最终希望找到极小值点或接近极小值点的点．这种算法的关键是要根据泛函指标 $J(u)$ 所提供的某些信息，确定一个能由第 i 点 u_i 求得第 $i+1$ 点 u_{i+1} 的搜索规则，抽象化之后，就是确定一个映 C 到 C 的搜索函数 $a=a(u)$，取 $u_{i+1}=a(u_i)$．更一般地，搜索函数不止一个，而是一个集合 $A=A(u)$，它是把 C 映入 C 内的点集的集值搜索函数，我们取 $u_{i+1}\in A(u_i)$．显然，搜索函数 $a=a(u)$ 是这里的特殊情形．于是，我们可建立下面的算法．

算法 1 （下降算法）

在 L^∞ 的某个闭子集 C 上确定集值搜索函数 $A(u)$．

第 0 步　取初始点 $u_0\in C$．

第 1 步　置 $i=0$．

第 2 步　取一点 $y\in A(u_i)$．

第 3 步　置 $u_{i+1}=y$．

第 4 步　若 $J(u_{i+1})\geqslant J(u_i)$，则停；否则，置 $i=i+1$ 并转第 2 步．

对于求解(UP)的上述算法，自然会提出点列 $\{u_i\}$ 能否收敛到(UP)的极小点上的问题．对这个问题往往不作直接的回答，而去证明点列 $\{u_i\}$ 的任一聚点都属于 L^∞ 内某个点集 Q，Q 的点或者是(UP)的极小点或者是满足某些最优性条件的点，两者必居其一．下面的定理就是解决 $\{u_i\}$ 的收敛性问题的．

定理 1　设泛函 $J(u)$ 在 L^∞ 内连续，C 是 L^∞ 内的一个子集．若对于 L^∞ 内任一点 $u\in C$ 但 $u\overline{\in} Q$，存在数 $\varepsilon(u)>0$ 和数 $\delta(u)<0$，致使对所有满足当 $\|u'-u\|_\infty\leqslant\varepsilon(u)$，$u'\in C$ 和 $u''\in A(u')$ 时有

$$J(u'')-J(u')\leqslant\delta(u),$$

则由算法 1 所构造出的点列 $\{u_i\}$，或者是有限序列 $\{u_0, u_1, \cdots,$

$u_k, u_{k+1}\}$，这时 $u_k \in Q$；或者 $\{u_i\}$ 是无限序列，这时它的每一个聚点 $u_* \in Q$。

证 首先，设 $\{u_i\}$ 为有限序列 $\{u_0, u_1, \cdots, u_k, u_{k+1}\}$. 这时必在算法1的第4步停止，即必定有 $J(u_{k+1}) \geqslant J(u_k)$，$u_{k+1} \in A(u_k)$，兹证，$u_k \in Q$，设不然，若 $u_k \overline{\in} Q$，则从假设条件知有

$$J(u_{k+1}) - J(u_k) \leqslant \delta(u_k) < 0,$$

得到矛盾，故必 $u_k \in Q$.

其次，设 $\{u_i\}$ 是无限的，u_* 是它的一个聚点，显然 $\{J(u_i)\}$ 是下降序列且有一收敛于 u_* 的子序列 $\{u_i\}$，$i \in K \subset \{0, 1, 2, \cdots\}$. 现证 $u_* \in Q$. 设不然 $u_* \overline{\in} Q$，则从假设条件知，存在 $\varepsilon(u_*) > 0$，$\delta(u_*) < 0$，和 $N \in K$，使对所有的 $i \in K$，$i \geqslant N$，有

$$\|u_i - u_*\|_\infty \leqslant \varepsilon(u_*),$$

和

$$J(u_{i+1}) - J(u_i) \leqslant \delta(u_*) < 0, \tag{5}$$

此处 $u_{i+1} \in A(u_i)$. 现在，任取子序列 $\{u_i\}$，$i \in K$，$i \geqslant N$，中的任意两个连接着的点 u_i, u_{i+j}，$j > 0$，注意到 $\{J(u_i)\}$ 是下降序列，又由式(5)就得

$$\begin{aligned} J(u_{i+j}) - J(u_i) &= [J(u_{i+j}) - J(u_{i+j-1})] + [J(u_{i+j-1}) \\ &\quad - J(u_{i+j-2})] + \cdots + [J(u_{i+1}) - J(u_i)] \\ &\leqslant J(u_{i+1}) - J(u_i) \leqslant \delta(u_*) < 0. \end{aligned}$$

由于 $J(u)$ 连续及 $u_i \to u_*$，$i \in K$，故先固定 $i \in K$，令 $i+j \to +\infty$，$i+j \in K$，再令 $i \to +\infty$，$i \in K$，从上式就得

$$\begin{aligned} 0 &= \lim_{i\to\infty} [J(u_*) - J(u_i)] \\ &= \lim_{i\to\infty} \lim_{j\to\infty} [J(u_{i+j}) - J(u_i)] \leqslant \delta(u_*) < 0, \end{aligned}$$

所得矛盾表明，$u_* \in Q$. ∎

注 这定理并未涉及到 $\{u_i\}$ 是否存在聚点的问题，而只证明

了若 $\{u_i\}$ 有聚点，则其任一聚点均属于 Q.

算法1的关键是选择搜索方向，即确定搜索规则 $A(u)$. 从 u_i 出发，人们总是希望在函数值 $J(u_i)$ 下降最快的那个方向去选取下一个点 u_{i+1}. 那么，哪个方向是函数下降最快的呢？在有限维空间内，例如设在三维空间内讨论函数 $F(z)$, $z=(x, y)\in\mathbb{R}^2$, 设 $h(z_i)$ 是过点 $z_i=(x_i, y_i)$ 的某个单位向量，则 $F(z)$ 过点 z_i 沿 $h(z_i)$ 的方向导数为：

$$\lim_{\varepsilon\to 0}\frac{F(z_i+\varepsilon h(z_i))-F(z_i)}{\varepsilon}$$

$$=\langle \operatorname{grad} F(z_i),\ h(z_i)\rangle$$

$$=\|\operatorname{grad} F(z_i)\|\|h(z_i)\|\cos\gamma=\|\operatorname{grad} F(z_i)\|\cos\gamma,$$

此处 $\operatorname{grad} F(z_i)=\left(\dfrac{\partial F(z_i)}{\partial x},\ \dfrac{\partial F(z_i)}{\partial y}\right)$ 是 $F(z)$ 过点 z_i 的梯度，γ 是 $h(z_i)$ 与 $\operatorname{grad} F(z_i)$ 之间的夹角. 显然，当 $\gamma=\pi$ 即 $h(z)$ 与 $\operatorname{grad} F(z_i)$ 反向时，$\langle \operatorname{grad} F(z_i),\ h(z_i)\rangle$ 取最小值，即负梯度方向是使函数 $F(z)$ 下降最快的方向.

对于无限维空间，也有上面这一结论，为此，我们把有限维空间的函数的梯度概念推广到无限维空间 $L^\infty([t_0, t_1], \mathbb{R}^m)$ 内.

定义 设 $G(u(\cdot))(\cdot)$: $L^\infty([t_0, t_1]; \mathbb{R}^m)\times[t_0, t_1]\to\mathbb{R}^m$, 对 $u(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)$ 有 $G(u(\cdot))\in L^\infty([t_0, t_1]; \mathbb{R}^m)$, 且对 $v(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)$ 有

$$\lim_{\|v\|_\infty\to 0}\frac{J(u+v)-J(u)-\langle G(u(\cdot)), v\rangle_\infty}{\|v\|_\infty}=0, \tag{6}$$

则称映象 $G(u(\cdot))(\cdot)$ 是 $J(\cdot)$ 的梯度，记为 $\operatorname{grad} J(\cdot)(\cdot)$ 或简记 $\operatorname{grad} J(\cdot)$.

梯度的几何意义 取 $v=\varepsilon w$, $\varepsilon>0$, $w\in L^\infty([t_0, t_1], \mathbb{R}^m)$, $\|w\|_\infty\neq 0$, 把 w 看作向量，令 ε 变化，则 v 就是沿方向 w 变化的向量，于是由式(6)可见

$$\lim_{\varepsilon\to0}\frac{J(u+v)-J(u)}{\varepsilon\|w\|_\infty}=\left\langle \operatorname{grad} J(u),\ \frac{w}{\|w\|_\infty}\right\rangle_\infty \tag{7}$$

从这式即可见，与有限维空间类似，可认为 $\left\langle \operatorname{grad} J(u),\ \frac{w}{\|w\|_\infty}\right\rangle_\infty$ 是在点 $v(\cdot)$ 上沿方向 w 的方向导数.

与有限维空间的情形一样，我们有

引理 1 若 $\operatorname{grad} J(\cdot)(\cdot)\neq0$，则负梯度 $-\operatorname{grad} J(u)$ 确定 $J(u)$ 的一个最速下降方向.

证 首先，在式(7)中取 $w=-\varepsilon \operatorname{grad} J(u)$，则得到

$$\lim_{\varepsilon\to0}\frac{J(u-\varepsilon \operatorname{grad} J(u))-J(u)}{\varepsilon}$$

$$=-\langle \operatorname{grad} J(u),\ \operatorname{grad} J(u)\rangle_\infty.$$

由此即得：存在数 $\varepsilon_0>0$，致使 $0\leqslant\varepsilon\leqslant\varepsilon_0$ 时有.

$$J(u-\varepsilon \operatorname{grad} J(u))-J(u)<0.$$

这个不等式说明了，当 $\operatorname{grad} J(u)\neq0$ 时，梯度确定了一个下降的方向 $-\operatorname{grad} J(u)$.

其次，从歌西-布尼亚柯夫斯基不等式有

$$|\langle \operatorname{grad} J(u),\ w\rangle_\infty|=|\langle \operatorname{grad} J(u),\ w\rangle_2|$$

$$\leqslant\|\operatorname{grad} J(u)\|_2\|w\|_2,$$

令

$$\langle \operatorname{grad} J(u),\ w\rangle_\infty=\|\operatorname{grad} J(u)\|_2\|w\|_2\rho,$$

则 $0\leqslant|\rho|\leqslant1$，故可写 $\rho=\cos\gamma$，即 $0\leqslant\gamma\leqslant\pi$，

$$\langle \operatorname{grad} J(u),\ w\rangle_\infty=\|\operatorname{grad} J(u)\|_2\|w\|_2\cos\gamma,$$

可把 γ 看作是 $\operatorname{grad} J(u)$ 与 w 之间的夹角. 于是，当且仅当 $\gamma=\pi$ 时，也即当且仅当 w 能使

$$\langle \operatorname{grad} J(u),\ w\rangle_\infty=-\|\operatorname{grad} J(u)\|_2\|w\|_2 \tag{8}$$

成立时，$\langle \operatorname{grad} J(u),\ w\rangle_\infty$ 达到最小值，而这不难验证，$w=-\varepsilon \operatorname{grad} J(u)$，$\varepsilon>0$，必使式(8)成立，故 $-\operatorname{grad} J(u)$ 是使 $J(u)$ 下降最速的方向. ∎

引理 2 若 $\hat{u}(\cdot)\in L^\infty([t_0, t_1]; \mathbb{R}^m)$ 使 $J(\hat{u})=\inf\{J(u)\mid u\in L^\infty([t_0, t_1]; \mathbb{R}^m)\}$，则必

$$\operatorname{grad} J(\hat{u})=0.$$

证 设不然，则从引理 1 知必存在 ε_0 使 $0<\varepsilon\leqslant\varepsilon_0$ 时

$$J(\hat{u}-\varepsilon \operatorname{grad} J(\hat{u}))-J(\hat{u})<0,$$

从而点 $\bar{u}=\hat{u}-\varepsilon\operatorname{grad} J(\hat{u})$ 使 $J(\bar{u})<J(\hat{u})$，这与 $\hat{u}$ 是 $J(u)$ 的极小值点相矛盾. ∎

总之，引理 1 表明，取负梯度方向作为求 $J(u)$ 的极小值点的方向，有可能加快整个计算过程；而引理 2 表明，为求 $J(u)$ 的极小值点，可取上面的定理 1 中的集 $Q=\{u\mid \operatorname{grad} J(u)=0,\ u\in L^\infty([t_0, t_1]; \mathbb{R}^m)\}$，及可用 $\operatorname{grad} J(u)$ 是否为零作为判断计算过程是否停止的法则. 在算法 1 中取负梯度方向作为搜索方向，就得如下的

算法 2 (最速下降算法或梯度法)

第 0 步 取初始点 $u_0\in C$.

第 1 步 置 $i=0$.

第 2 步 计算 $\operatorname{grad} J(u_i)$.

第 3 步 置 $h(u_i)=-\operatorname{grad} J(u_i)$. 若 $h(u_i)=0$，则停止计算；否则，至第 4 步.

第 4 步 计算数 $\lambda(u_i)$，它是满足

$$J(u_i+\lambda(u_i)h(u_i))=\inf\{J(u_i+\lambda h(u_i))\mid\lambda\geqslant 0\}$$

的最小的非负数.

第 5 步 置 $u_{i+1}=u_i+\lambda(u_i)h(u_i)$，置 $i=i+1$，转第 2 步.

这样就把高维极小问题转化为求 $J(u+\lambda h(u))$ 关于 $\lambda\geqslant 0$ 的一维极小值问题了. 在算法 2 中的第 4 和第 5 步是不能直接由数字计算机来执行的，于是自然会提出怎样进行这种形式的一维搜索的问题. 对此，人们采取了各种方法来修改算法 2，找出

$J(u_i+\lambda h(u_i))$的极小值点的一个近似值 u_{i+1}，而使产生的 $\{u_i\}$ 仍保持收敛性.

下面介绍由 Goldstein 提出的梯度算法. 其主要思想可通过下面的几何意义来说明.

给曲线

$$\theta(\lambda)=J(u+\lambda h(u))-J(u), \quad \lambda\geqslant 0,$$

其中 $\theta'(0)=\langle \operatorname{grad} J(u), h(u)\rangle$，已知 $\theta'(0)<0$，故 $\theta(\lambda)$在点 $\lambda=0$ 附近是下降的. 我们的目的是寻找一个大于 $\lambda=0$ 的 $\theta(\lambda)$ 的近似极小值点，我们利用导数 $\theta'(\lambda)=0$ 的信息作为终止准则. 这样的 λ 有可能使 $\theta(\lambda)$最小，但也不一定，如图 5.1，由于 $\theta(c)<\theta(0)$且 $\theta'(c)=0$，故若试探点取到点 $\lambda=c$ 时，一维搜索会停止，但从 $\theta(0)$ 到 $\theta(c)$ 的下降量却很小. 为预防这种情形的发生，我们从原点出发作射线

$$l_1:\ y=\alpha\theta'(0)\lambda=\lambda\alpha\langle \operatorname{grad} J(u),\ h(u)\rangle,$$

此处 $0<\alpha<\frac{1}{2}$，限制在射线 l_1 的下方来求曲线 $y=\theta(\lambda)$ 的近似极小值点. 但是，若只做上述限制还是不妥的，从图 5.1 可见，整个区间$(0, b]$内的 λ 都使 $\theta(\lambda)\leqslant\alpha\theta'(0)\lambda$，因而不论多么接近 0 的 λ 都

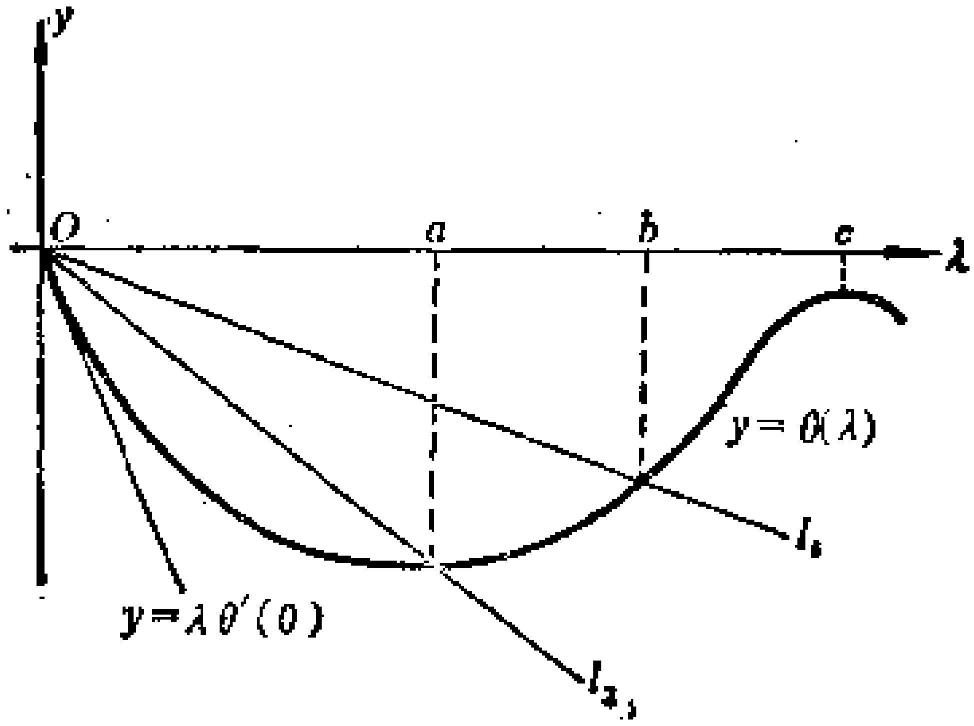


图 5.1

是$\theta(\lambda)$的可接受的近似极小值点，而这可能使一维搜索求得的点λ与初始点$\lambda=0$没有多大差别，因此再取如下的一个对λ的限制条件，作过原点的射线

$$l_2:\ y=(1-\alpha)\theta'(0)\lambda=\lambda(1-\alpha)\langle \operatorname{grad} J(u),\ h(u)\rangle,$$

考虑在射线l_2的上方求曲线$y=\theta(\lambda)$的近似极小点．这样，当$\lambda\in[a,\ b]$时，曲线$y=\theta(\lambda)$在射线l_1和l_2之间，此时函数值$\theta(\lambda)$有所下降(对数值$\theta(0)$而言)，并且$|\theta'(\lambda)|$相当小，因此，$\lambda\in[a,\ b]$是可接受的关于极小值问题$\min\{\theta(\lambda)\,|\,\lambda\geqslant 0\}$的近似极小值点．$l_1$和$l_2$就成为确定$\lambda$的两个限制准则．选取$\alpha\in\left(0,\ \frac{1}{2}\right)$且靠近$\frac{1}{2}$，经研究会使算法中的序列$\{u_i\}$有较快的收敛速度．本着这种想法，就可提出下面修改了的梯度法．

算法 3 (Goldstein 梯度法)

第 0 步　选取初始点$u_0\in L^\infty([t_0,\ t_1];\ \mathbb{R}^m)$使集

$$C(u_0)\triangleq\{u\,|\,J(u)\leqslant J(u_0),\ u\in L^\infty([t_0,\ t_1];\ \mathbb{R}^m)\}$$

有界；选取$\alpha\in\left(0,\ \frac{1}{2}\right)$.

第 1 步　置$i=0$.

第 2 步　计算$\operatorname{grad} J(u_i)$.

第 3 步　置$h(u_i)=-\operatorname{grad} J(u_i)$．若$h(u_i)=0$，则停止计算；否则，至第 4 步．

第 4 步　计算数$\lambda_i=\lambda(u_i)$，它使

$$\lambda_i(1-\alpha)\frac{\partial}{\partial\lambda}\theta(0,\ u_i)\leqslant\theta(\lambda_i,\ u_i)\leqslant\lambda_i\alpha\frac{\partial}{\partial\lambda}\theta(0,\ u_i),$$

此处
$$\theta(\lambda,\ u_i)=J(u_i+\lambda h(u_i))-J(u_i),$$

$$\frac{\partial}{\partial\lambda}\theta(0,\ u_i)=\langle \operatorname{grad} J(u_i),\ h(u_i)\rangle_\infty.$$

第 5 步　置$u_{i+1}=u_i+\lambda(u_i)h(u_i)$，置$i=i+1$，转第 2 步．

上述算法中的梯度与 $x(\cdot, u)$ 有关，这是最优控制问题的特点和难点. 为了体现此特点和更切实用，我们将根据所研究的最优控制问题(1)、(2)计算出 $\operatorname{grad} J(u)$ 的近似值.

对容许控制 $v(\cdot)$，如第四章那样，补充引进坐标 $x^0(\cdot, v)$，适合

$$\frac{dx^0(t, v)}{dt}=f^0(t, x(t, v), v(t)),\ x^0(t_0)=0,$$

那末

$$x^0(t_1, v)=J(v(\cdot)).$$

现取定容许控制 $u(\cdot)$ 和任定的 $\delta u(\cdot)\in L^\infty([t_0, t_1]; R^m)$，则 $u^\varepsilon(\cdot)=u(\cdot)+\varepsilon\delta u(\cdot)$ 仍然是容许的. 置 $\boldsymbol{x}^\varepsilon(\cdot)=\boldsymbol{x}^\varepsilon(\cdot, u^\varepsilon)=(x^{0\varepsilon}(\cdot), x^\varepsilon(\cdot))$，注意到 $\boldsymbol{f}=(f^0, f)^T$ 关于 u 是连续的，就可推得定理 IV-2-1 的结论仍成立，即

$$\boldsymbol{x}^\varepsilon(t)=\boldsymbol{x}(t)+\varepsilon\delta\boldsymbol{x}(t)+o(\varepsilon),$$

此处 $\boldsymbol{x}(t)=\boldsymbol{x}(t, u(\cdot))$，而 $\delta\boldsymbol{x}(t)=(\delta x^0(t), \delta x(t))^T$ 则适合变分微分方程组

$$\frac{d}{dt}\delta\boldsymbol{x}(t)=\left(\frac{\partial\boldsymbol{f}(t, x(t), u(t))}{\partial x}\right)^T\delta x(t)+\left(\frac{\partial\boldsymbol{f}(t, x(t), u(t))}{\partial u}\right)^T\delta u(t).$$

这样，从上面诸关系式可见

$$\begin{aligned}
&J(u^\varepsilon(\cdot))-J(u(\cdot))\\
&\quad=x^{0\varepsilon}(t_1)-x^0(t_1)\\
&\quad=\varepsilon\int_{t_0}^{t_1}\left[\left(\frac{\partial f^0(t, x(t), u(t))}{\partial x}\right)^T\delta x(t)\right.\\
&\qquad\left.+\left(\frac{\partial f^0(t, x(t), u(t))}{\partial u}\right)^T\delta u(t)\right]dt+o(\varepsilon),
\end{aligned}$$

此处 $\delta x(t)$ 满足线性微分方程组

$$\begin{cases} \dfrac{d}{dt}\delta x(t) = \left(\dfrac{\partial f(t, x(t), u(t))}{\partial x}\right)^T \delta x(t) \\ \qquad\qquad + \left(\dfrac{\partial f(t, x(t), u(t))}{\partial u}\right)^T \delta u(t), \\ \delta x(t_0) = 0. \end{cases} \tag{9}$$

根据微分方程理论(参见第一章 § 5),方程(9)的解是

$$\delta x(t) = \int_{t_0}^{t} \Phi(t, s)\left[\frac{\partial f(s, x(s), u(s))}{\partial u}\right]^T \delta u(s)ds,\ t \in [t_0, t_1],$$

此处 $\Phi(t, s)$ 是对应于(9)的齐次线性微分方程组

$$\frac{dy}{dt} = \left(\frac{\partial f(t, x(t), u(t))}{\partial x}\right)^T y$$

的基本解方阵, $\Phi(s, s) = I$, I 是 $n\times n$ 单位矩阵. 这样, 即得

$$\begin{aligned} & J(u^\varepsilon(\cdot)) - J(u(\cdot)) \\ &= \varepsilon \int_{t_0}^{t_1}\int_{t_0}^{t}\left(\frac{\partial f^0(t, x(t), u(t))}{\partial x}\right)^T \Phi(t, s) \\ &\quad \times \left(\frac{\partial f(s, x(s), u(s))}{\partial u}\right)^T \delta u(s)ds\,dt \\ &\quad + \varepsilon\int_{t_0}^{t_1}\left(\frac{\partial f^0(t, x(t), u(t))}{\partial u}\right)^T \delta u(t)dt + o(\varepsilon) \\ &= \varepsilon\int_{t_0}^{t_1}\left[\int_{s}^{t_1}\left(\frac{\partial f^0(t, x(t), u(t))}{\partial x}\right)^T \Phi(t, s)dt\right] \\ &\quad \times \left(\frac{\partial f(s, x(s), u(s))}{\partial u}\right)^T \delta u(s)ds \\ &\quad + \varepsilon\int_{t_0}^{t_1}\left(\frac{\partial f^0(s, x(s), u(s))}{\partial u}\right)^T \delta u(s)ds + o(\varepsilon); \end{aligned} \tag{10}$$

我们在计算右边第一项时应用了 Dirichlet 换限公式

$$\int_a^b dt\int_a^t g(t, s)ds = \int_a^b ds\int_s^b g(t, s)dt$$

而交换了积分顺序.

现在考虑微分方程组

$$\frac{d}{ds}\psi(s)=-\frac{\partial f(s,x(s),u(s))}{\partial x}\psi(s)+\frac{\partial f^0(s,x(s),u(s))}{\partial x},\tag{11}$$

求它的满足初值 $\psi(t_1)=0$ 的特解，则由微分方程组的理论(参见第一章)知

$$\psi(s)=\int_{t_1}^{s}\Phi^T(t,s)\frac{\partial f^0(t,x(t),u(t))}{\partial x}dt,\qquad s\in[t_0,t_1],$$

此处 $\Phi^T(t,s)$ 是 $\Phi(t,s)$ 的转置方阵，代入(10)即得

$$\begin{aligned}
&J(u^\varepsilon(\cdot))-J(u(\cdot))\\
&=-\varepsilon\int_{t_0}^{t_1}\psi^T(s)\left(\frac{\partial f(s,x(s),u(s))}{\partial u}\right)^T\delta u(s)ds\\
&\quad+\varepsilon\int_{t_0}^{t_1}\left(\frac{\partial f^0(s,x(s),u(s))}{\partial u}\right)^T\delta u(s)ds+o(\varepsilon)\\
&=-\varepsilon\int_{t_0}^{t_1}\left\langle\frac{\partial f(s,x(s),u(s))}{\partial u}\psi(s)\right.\\
&\quad\left.+\frac{\partial f^0(s,x(s),u(s))}{\partial u},\ \delta u(s)\right\rangle ds+o(\varepsilon).
\end{aligned}\tag{12}$$

另一方面，从梯度的定义出发，在(7)中取 $w=\delta u$，则得

$$\begin{aligned}
&J(u^\varepsilon(\cdot))-J(u(\cdot))\\
&=\varepsilon\langle\operatorname{grad}J(u(\cdot)),\ \delta u(\cdot)\rangle_\infty+o(\varepsilon),
\end{aligned}$$

将此式与(12)式相比较就得到

$$\begin{aligned}
\operatorname{grad}J(u)(t)=&-\frac{\partial f(t,x(t),u(t))}{\partial u}\psi(t)\\
&+\frac{\partial f^0(t,x(t),u(t))}{\partial u}.
\end{aligned}\tag{13}$$

注意，此处 $\psi(t)$ 是方程(11)的初值为 $\psi(t_1)=0$ 的解。

现在，我们可以把求解最优化问题(UP)即最优控制问题(1)、(2)的算法 3 改写成下面的

算法 4 (Goldstein 梯度法)

第 0 步　选取一个初始控制 $u_0(\cdot)\in L^\infty([t_0,t_1],\mathbb{R}^m)$，使(9)

的集 $C(u_0)$ 有界；选取 $\alpha\in\left(0,\ \frac{1}{2}\right)$.

第 1 步　置 $i=0$.

第 2 步　取 $u(t)=u_i(t)$, $t\in[t_0,\ t_1]$, 解微分方程(1)而得

$$x_i(t)=x(t,\ u_i),\ t\in[t_0,\ t_1].$$

第 3 步　在 $u(t)=u_i(t)$, $x(t)=x_i(t)$, $t\in[t_0,\ t_1]$, 之下求方程(11)满足 $\psi_i(t_1)=0$ 的特解 $\psi_i(t)$, $t\in[t_0,\ t_1]$.

第 4 步　按式(13)计算

$$\operatorname{grad} J(u_i)(t)=-\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t)+\frac{\partial f^0(t,\ x_i(t),\ u_i(t))}{\partial u},\quad t\in[t_0,\ t_1].$$

第 5 步　置 $h(u_i)(t)=-\operatorname{grad} J(u_i)(t)$, 若 $h(u_i)(\cdot)=0$, 则停；否则，到第 6 步.

第 6 步　计算数 $\lambda_i=\lambda(u_i(\cdot))>0$ 致使

$$-\lambda_i(1-\alpha)\|\operatorname{grad} J(u_i)\|_2^2\leqslant\theta(\lambda_i,\ u_i)\leqslant-\lambda_i\alpha\|\operatorname{grad} J(u_i)\|_2^2,$$

此处
$$\theta(\lambda_i,\ u_i)=J(u_i+\lambda_i h(u_i))-J(u_i).$$

第 7 步　置 $u_{i+1}(t)=u_i(t)+\lambda_i(u_i)h(u_i)(t)$, $t\in[t_0,\ t_1]$; 置 $i=i+1$, 且转第 2 步.

我们还应讨论，由算法 4 所构造的 $\{u_i\}$ 的收敛性. 下面的定理回答了这个问题.

定理 2　设控制序列 $\{u_i(\cdot)\}$ 由算法 4 产生，则

1°　若 $\{u_i(\cdot)\}$ 是有限序列 $\{u_i(\cdot)\}=\{u_1(\cdot),\ \cdots,\ u_k(\cdot)\}$, k 是某整数，则 $\operatorname{grad} J(u_k)(\cdot)=0$.

2°　若 $\{u_i(\cdot)\}$ 是无限序列，则它的每一个聚点 $u_*(\cdot)$ 必满足

$$\operatorname{grad} J(u_*)(\cdot)=0.$$

证　我们将指出算法 4 是算法 1 的特殊情形，并且满足定理 1 中的假设条件. 取 $C\triangleq C(u_0)$ 及取 $Q\triangleq\{u(\cdot)\,|\,u(\cdot)\in$

$L^\infty([t_0,t_1],\mathbb{R}^m)$, $\mathrm{grad}\,J(\bar{u})(\cdot)=0\}$，并在空间 $L^\infty \triangleq L^\infty([t_0, t_1]; \mathbb{R}^m)$ 内按下式定义集值搜索函数 $A(u)$：

$$A(u(\cdot))=\begin{cases} u, & \text{当 } u\in Q, \\ \{\bar{u}(\cdot)=u(\cdot)+\lambda h(u)(\cdot) \mid \lambda\geqslant 0, \\ \qquad \underline{\theta}(\lambda, u)\geqslant 0, \bar{\theta}(\lambda, u)\leqslant 0\}, & \text{当 } u\overline{\in} Q, \end{cases}$$

此处 $h(u)(\cdot)$ 是算法 4 的第 5 步中的映象(删去了下标)，即

$$h(u)(t)\triangleq -\mathrm{grad}\,J(u)(t),$$

而

$$\begin{aligned} \underline{\theta}(\lambda, u) &\triangleq \theta(\lambda, u)-[-\lambda(1-\alpha)\|\mathrm{grad}\,J(u)\|_2^2] \\ &= [J(u+\lambda h(u))-J(u)]+\lambda(1-\alpha)\|\mathrm{grad}\,J(u)\|_2^2, \\ \bar{\theta}(\lambda, u) &\triangleq \theta(\lambda, u)-[-\lambda\alpha\|\mathrm{grad}\,J(u)\|_2^2 \\ &= [J(u+\lambda h(u))-J(u)]+\lambda\alpha\|\mathrm{grad}\,J(u)\|_2^2. \end{aligned}$$

显然，$\lambda\geqslant 0$, $\underline{\theta}(\lambda, u)\geqslant 0$, $\bar{\theta}(\lambda, u)\leqslant 0$ 等价于算法 4 第 6 步中的不等式．现在已易看出算法 4 是算法 1 的特殊情形了．定理的结论 1° 是显见的，故剩下的去指出结论 2° 成立即可．

首先，易见 $J(u)(\cdot)$ 是连续的，下面我们将指出，映象 $J(u)(\cdot)$ 和 $A(\cdot)$ 满足定理 1 的其余的假设条件．作函数

$$\varphi(\lambda, u)=\frac{\underline{\theta}(\lambda, u)}{\lambda},$$

规定 $\varphi(0, u)=\lim\limits_{\lambda\to 0}\varphi(\lambda,u)$，则 $\varphi(\lambda,u):[0,+\infty)\times L^\infty([t_0,t_1];\mathbb{R}^m)\to\mathbb{R}$ 是连续映象．任取定点 $u(\cdot)\in C(u_0)$ 但 $u(\cdot)\overline{\in} Q$，易见 $\varphi(0, u)<0$；另方面，因集 $C(u_0)$ 是有界的，故过点 u 的线段 $u+\lambda h(u)$, $\lambda\geqslant 0$, 必有点 $\overline{\in} C(u_0)$，即存在 $\bar{\lambda}>0$ 使 $J(u+\bar{\lambda}h(u))>J(u)$，从而 $\varphi(\bar{\lambda}, u)>0$．对给定的 $u(\cdot)$, $\varphi(\lambda, u)$ 是单参数 $\lambda\in[0, \bar{\lambda}]$ 的连续函数，且注意到 $\underline{\theta}(\lambda, u)$ 的形式，就知必存在最小的正零点 $\rho(u)>0$，使

$$\varphi(\lambda, \mu)=0, \quad \text{当 } \lambda=\rho(u),$$

$$\varphi(\lambda,\ \mu)<0,\quad 当\ \lambda\in[0,\ \rho(u));$$

且当 $0\leqslant\lambda\leqslant\rho(u)$ 时，点 $\bar{u}=u+\lambda h(u)\in O(u_0)$.

$$\gamma(u)=\max\left\{\varphi(\lambda,\ u)\,|\,\lambda\in\left[0,\ \frac{\rho(u)}{2}\right]\right\}<0. \tag{14}$$

现在验证定理 1 的条件确实成立. 取 $u\in O(u_0)$ 但 $u\overline{\in}Q$. 从上面知存在数 $\gamma(u)<0$, 于是, 因 $\varphi(\lambda,\ u)$ 是连续函数, 故存在 $\varepsilon_1(u)>0$ 使 $u'\in\{u'\,|\,\|u'-u\|_\infty\leqslant\varepsilon_1\}$ 和 $\lambda\in[0,\ \rho(u)/2]$ 时有

$$|\varphi(\lambda,\ u')-\varphi(\lambda,\ u)|\leqslant-\frac{\gamma(u)}{2};$$

从而注意到式(14)就得: 当 $\lambda\in\left[0,\ \frac{\rho(u)}{2}\right]$, $u'\in\{u'\,|\,\|u'-u\|_\infty\leqslant\varepsilon_1\}$ 时

$$\varphi(\lambda,\ u')\leqslant\frac{1}{2}\gamma(u). \tag{15}$$

由此我们得到下面的断语: 对任一个 $u'\in\{u'\,|\,\|u'-u\|_\infty\leqslant\varepsilon_1\}$, 若取 $u''\in A(u')$ 则必有

$$u''=u'+\lambda'h(u'),\qquad \lambda'>\frac{\rho(u)}{2}. \tag{16}$$

事实上，设不然，若 $0<\lambda'\leqslant\rho(u)/2$, 则从(15)得

$$\frac{\underline{\theta}(\lambda',\ u')}{\lambda'}=\varphi(\lambda',\ u')\leqslant\frac{1}{2}\gamma(u)<0;$$

但 $u''\in A(u')$ 从而 $\underline{\theta}(\lambda',\ u')\geqslant0$, 所得矛盾证明了式(16)成立.

我们从

$$\|\mathrm{grad}\,J(u)\|_2^2=\int_{t_0}^{t_1}\langle\mathrm{grad}\,J(u),\ \mathrm{grad}\,J(u)\rangle dt$$

及式(13), 注意到 $\frac{\partial f}{\partial u}$, $\frac{\partial f^0}{\partial u}$ 关于 t, x, u 的连续性, 就可见 $\|\mathrm{grad}\,J(u)\|_2^2$ 关于 $u(\cdot)\in L^\infty$ 连续; 又因 $u\in O(u_0)$ 和 $u\overline{\in}Q$, 故对正数 $\frac{1}{2}\|\mathrm{grad}\,J(u)\|_2^2>0$, 存在 $\varepsilon_2(u)>0$, 使对任一个 $u'\in\{u'\,|\,\|u'-u\|_\infty\leqslant\varepsilon_2\}$ 有

$$0<\frac{1}{2}\|\text{grad}\,J(u)\|_2^2\leqslant\|\text{grad}\,J(u')\|_2^2. \tag{17}$$

取 $\varepsilon(u)=\min\{\varepsilon_1,\ \varepsilon_2\}$，于是，对所有的

$$u'\in\{u'\,|\,\|u'-u\|_\infty\leqslant\varepsilon(u)\},\ u''=u'+\lambda'h(u')\in A(u'),$$

从 $\bar{\theta}(\lambda',\ u')\leqslant 0$，(16)和(17)必得

$$J(u'')-J(u')\leqslant-\lambda'\alpha\|\text{grad}\,J(u')\|_2^2$$

$$\leqslant-\frac{1}{4}\rho(u)\alpha\|\text{grad}\,J(u)\|_2^2<0.$$

置 $\delta(u)\triangleq-\frac{1}{4}\rho(u)\alpha\|\text{grad}\,J(u)\|_2^2$，即得证对泛函 $J(u)$ 来说，定理 1 的假设条件成立。这样，应用定理 1，本定理的结论 2° 也成立。

推论 考虑集 $Q=\{u\,|\,\text{grad}\,J(u)=0,\ u\in L^\infty([t_0,\ t_1];\ \mathbb{R}^m)\}$，且假设对 Q 内的任何 $u'\neq u''$ 均有 $J(u')\neq J(u'')$。若集 $C(u_0)=\{u\,|\,J(u)\leqslant J(u_0)\}$ 是紧致的，则由算法 4 所产生的无穷序列 $\{u_i\}$ 必定收敛于唯一的一点 $u_*\in Q$。

证 因为 $C(u_0)$ 紧致，故 $\{u_i\}$ 必有聚点。其次，由算法 4 可见 $\{J(u_i)\}$ 单调下降，因 $C(u_0)$ 紧致和 $J(u)$ 连续，故必有

$$J(u_i)\to J_*>-\infty,\ \text{当}\ i\to+\infty.$$

设 u', u'' 都是 $\{u_i\}$ 的聚点，且 $u'\neq u''$，则由假设知 $J(u')\neq J(u'')$。另方面，从上面定理 2 可见 u'、$u''\in Q$，且因为 $J(u)$ 连续和 $\{J(u_i)\}$ 收敛于 J_*，故必有

$$J(u')=J_*=J(u''),$$

这与假设相悖。由此矛盾得证推论成立。 ∎

习　题

1　给出线性二次型最优控制问题，状态方程和性能指标为

$$\begin{cases} \dfrac{d}{dt}x(t)=A(t)x(t)+B(t)u(t), & t\in[t_0,\ t_1], \\ x(t_0)=x_0 \end{cases}$$

$$J(u(\cdot))=\frac{1}{2}\int_{t_0}^{t_1}[\langle x(t),\ R(t)x(t)\rangle+\langle u(t),\ Q(t)u(t)\rangle]dt,$$

此处 $x(t)\in\mathbb{R}^n$, $u(t)\in\mathbb{R}^m$; $R(t)$ 和 $Q(t)$ 是对称正定矩阵, $R(\cdot)$, $Q(\cdot)$, $A(\cdot)$, $B(\cdot)$均连续. 试建立相应于算法 4 的算法, 并证明由算法所提供的控制序列 $\{u_i\}$ 收敛到二次型最优控制问题的一个最优控制.

2 求证满足算法 4 第 6 步中的不等式的数 λ 必存在.

3 试设计出一个计算满足算法 4 第 6 步的不等式的数 λ_i 的算法.

§3 牛顿法与边值问题

庞特里雅金最大值原理是最优控制问题的必要条件, 从计算方法的角度来看, 这组必要条件是一种特殊类型的边值问题, 故可用求解微分方程边值问题的各种方法来求解最优控制问题. 本节将介绍如何用 Newton 法来解最优控制问题.

我们先考察求通常的单变量 x 的方程

$$g(x)=0 \tag{1}$$

之根的问题. Newton 法的基本思想是: 把求解方程(1)的问题转化为求

$$\begin{cases} y=g(x), & (2)\\ y=0 & (3)\end{cases}$$

之交点的问题. 对曲线 $y=g(x)$ 用一条直线去近似, 然后求出此直线与直线 $y=0$ 的交点, 以它作为(2)和(3)之交点的近似值, 具体地说, 设已知方程(2)和(3)的解的第 0 次近似 x_0, 则用曲线 $y=g(x)$在$(x_0,\ g(x_0))$上的切线

$$y=g(x_0)+g'(x_0)(x-x_0) \tag{4}$$

来作曲线(2)的近似, 于是, 求直线(4)与直线(3)的交点 x_1 作为曲线(2)和直线(3)的第 1 次近似, 即 x_1 满足

$$g(x_0)+g'(x_0)(x-x_0)=0.$$

重复使用这一近似方法，就得到第 $i+1$ 次近似 x_{i+1} 满足方程式

$$g(x_i)+g'(x_i)(x_{i+1}-x_i)=0;$$

设 $(g'(x_i))^{-1}$ 存在，则得

$$x_{i+1}=x_i-[g'(x_i)]^{-1}g(x_i). \tag{5}$$

因此，求解方程(1)的问题，就转化为按公式(5)求近似序列 $\{x_i\}$ 的问题了.

上述思想和方法同样可推广到求解如下的函数方程

$$G(z)=0, \tag{6}$$

此处 G: $B\to B$, B 是 Banach 空间. 定义 $G(\cdot)$在点 $z\in B$ 上的导数$\left(\frac{\partial G(z)}{\partial z}\right)(\cdot)$为映 B 到 B 的线性泛函，它满足

$$\lim_{\|\delta z\|_B\to 0}\frac{\left\langle G(z+\delta z)-G(z)-\frac{\partial G(z)}{\partial z}(\delta z)\right\rangle_B}{\delta z}=0, \tag{7}$$

此处 $\|\cdot\|_B$ 表在空间 B 的模，$\delta G(z)\triangleq\frac{\partial G(z)}{\partial z}(\delta z)$ 是 $G(z)$ 在点 z 的增量 $\Delta G(z)=G(z+\delta z)-G(z)$ 关于 δz 的线性主部. 设$\left(\frac{\partial G(z)}{\partial z}\right)(\cdot)$ 及其逆映象$\left(\frac{\partial G(z)}{\partial z}\right)^{-1}(\cdot)$存在且连续，与上面求有限维方程(1)的根完全类似，为解函数方程(6)，可应用 Newton 法. 即把 $G(z)$在 i 次近似 z_i 上展开，取线性部分：

$$G(z)\doteq G(z_i)+\frac{\partial G(z_i)}{\partial z}(z-z_i); \tag{8}$$

方程(6)的解的第 $i+1$ 次近似 z_{i+1} 认为是使 $G(z_{i+1})=0$，于是得到 z_{i+1} 满足

$$G(z_i)+\frac{\partial G(z_i)}{\partial z}(z_{i+1}-z_i)=0. \tag{9}$$

用 $z=z_{i+1}$ 代入(8)，记 $\delta z_i=z_{i+1}-z_i$，也可把(8)和(9)写成

$$\Delta G(z_i) = G(z_i + \delta z_i) - G(z_i) \doteq \delta G(z_i), \tag{10}$$

$$G(z_i) + \delta G(z_i) = 0. \tag{11}$$

因此，为了求方程(6)的第 $i+1$ 次近似解 z_{i+1}，首先求泛函 $G(z)$ 在点 z 上的增量 $\Delta G(z_i)$ 的线性主部 $\delta G(z_i)$，用它代替 $\Delta G(z_i)$，这就是式(10)；然后令式(11)成立而求出点 z_{i+1}，可以认为它使 $G(z)$ 近似地等于零，从而得 z_{i+1} 是方程(6)的近似解．在 $G(z)$ 满足一定的条件下，按式(9)或即(11)构造出的序列 $\{z_i\}$ 是收敛的(参见关肇直《泛函分析讲义》[4])，设其极根点是 z 则由于已设 $G(\cdot)$，$\left(\frac{\partial G(\cdot)}{\partial z}\right)(\cdot)$ 连续，从式(11)就得 $G(z_*)=0$，从而获得方程(6)的解，由上述可见，为求函数方程(6)的解，Newton 法可采取下面的形式．

算法 1 (Newton 法)

第 0 步　选取一点 $z_0 \in B$.

第 1 步　置 $i=0$.

第 2 步　计算 $G(z_i)$.

第 3 步　若 $G(z_i)=0$ 停；否则至第 4 步.

第 4 步　按公式(11)计算 z_{i+1}.

第 5 步　置 $i=i+1$，返回第 2 步.

下面我们将应用上述的在 Banach 空间内讨论的 Newton 法来求解最优控制问题.

现在设受控过程的状态方程是

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad t \in [t_0, t_1], \tag{12}$$

$$x(t_0) = x_0, \quad g(x(t_1)) = 0. \tag{13}$$

性能指标为

$$J(u(\cdot)) = \int_{t_0}^{t_1} f^0(t, x(t), u(t))dt, \tag{14}$$

此处 $f^0:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $g:\mathbb{R}^n\to\mathbb{R}^s$. $(s\leqslant n)$, 关于 x 和 u 连续可微; $f^0, b, \dfrac{\partial f^0}{\partial x}, \dfrac{\partial f^0}{\partial u}, \dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial u}$ 关于 t 连续; 还设 $\dfrac{\partial g}{\partial x}$ 在 $\mathbb{R}^n$ 内有最大秩; t_0、t_1 和 x_0 已给定, 取控制 $u(\cdot)\in L^\infty\triangleq L^\infty([t_0, t_1];\mathbb{R}^m)$的有界可测函数的等价类空间.

最优控制问题 选取容许控制 $u(\cdot)\in L^\infty$, 使相应的系统(12)在 t_0 时刻从 x_0 出发满足终端条件 $g(x(t_1))=0$ 的轨线 $x(t)=x(t, u)$ 致使

$$J(u(\cdot))=\inf\{J(v(\cdot)\,|\,x(t_0, v)=x_0,\quad g(x(t_1, v))=0,\ v(\cdot)\in L^\infty\}.$$

根据最大值原理得知, 若 $u(\cdot)$ 是最优控制, $x(\cdot)$ 是相应的最优轨线, 则必存在 $\psi_0=-1$(即我们这里假定为非退化的情形), $\psi(\cdot)=(\psi_1(\cdot), \cdots, \psi_n(\cdot))^T$ 满足方程

$$\frac{d}{dt}\psi(t)=\frac{\partial f^0(t, x(t), u(t))}{\partial x}-\frac{\partial f(t, x(t), u(t))}{\partial x}\psi(t).$$

又根据横截条件, 注意到端点条件(13)的形式知, 存在向量 $e=(e_1, e_2, \cdots, e_s)^T$ 使

$$\psi(t_1)=\frac{\partial g(x(t_1))}{\partial x}e.$$

而最大值条件是: 在 $[t_0, t_1]$ 上几乎处处有

$$\begin{aligned}&\mathscr{H}(t, \psi(t), x(t), u(t))\\&=\max_{v\in\mathbb{R}^m}\mathscr{H}(t, \psi(t), x(t), v),\end{aligned}\tag{15}$$

此处

$$\mathscr{H}(t, \psi, x, u)=-f^0(t, x, u)+\langle\psi, f(t, x, u)\rangle.$$

因此, $(x(\cdot), \psi(\cdot), u(\cdot), e)^T$ 必定是下述函数方程组的解: 在 $t\in[t_0, t_1]$ 上

$$x_0+\int_{t_0}^t f(\tau, x(\tau), u(\tau))d\tau-x(t)=0,\tag{16}$$

$$g(x(t_1))=0, \tag{17}$$

$$\frac{\partial g(x(t_1))}{\partial x}e+\int_{t_1}^{t}\left[\frac{\partial f^0(\tau, x(\tau), u(\tau))}{\partial x}\right.$$

$$\left.-\frac{\partial f(\tau, x(\tau), u(\tau))}{\partial x}\psi(\tau)\right]d\tau-\psi(t)=0, \tag{18}$$

$$-\frac{\partial f^0(t, x(t), u(t))}{\partial u}+\frac{\partial f(t, x(t), u(t))}{\partial u}\psi(t)=0. \tag{19}$$

后一个式子是从最大值条件(15)对 $\mathscr{H}(t, \psi, x, u)$ 关于 u 求异而推得的.

现在,我们把(16)—(19)表为方程组

$$G(z)=0$$

的形式,此处 $G:B\to B$, B 是 Banach 空间。为此,令 $B=C_n[t_0, t_1]\times C_n[t_0, t_1]\times L^\infty\times\mathbb{R}^s$, $C_n[t_0, t_1]$ 是 $[t_0, t_1]\to\mathbb{R}^n$ 的连续函数空间. 设 $z=(x(\cdot), \psi(\cdot), u(\cdot), e)^T$ 为 B 内的点,记

$$\|z\|_B=[(\sup_{t\in[t_0,t_1]}\|x(t)\|)^2+(\sup_{t\in[t_0,t_1]}\|\psi(t)\|)^2+\|u\|_\infty^2$$
$$+\|e\|^2]^{1/2},$$

则易验证 B 是 Banach 空间.

现在引入映象:在 $t_0\leqslant t\leqslant t_1$ 上

$$G_1(x(\cdot), \psi(\cdot), u(\cdot), e)(t)$$
$$=x_0+\int_{t_0}^{t}f(\tau, x(\tau), u(\tau))d\tau-x(t),$$

$$G_2(x(\cdot), \psi(\cdot), u(\cdot), e)(t)$$
$$=\frac{\partial g(x(t_1))}{\partial x}e+\int_{t_1}^{t}\left\{\frac{\partial f^0(\tau, x(\tau), u(\tau))}{\partial x}\right.$$
$$\left.-\frac{\partial f(\tau, x(\tau), u(\tau))}{\partial x}\psi(\tau)\right\}d\tau-\psi(t),$$

$$G_3(x(\cdot), \psi(\cdot), u(\cdot), e)(t)$$
$$=-\frac{\partial f^0(t, x(t), u(t))}{\partial u}+\frac{\partial f(t, x(t), u(t))}{\partial u}\psi(t).$$

定义

$$G(z)=G(x(\cdot),\ \psi(\cdot),\ u(\cdot),\ e)$$

$$=\begin{bmatrix} G_1(x(\cdot),\ \psi(\cdot),\ u(\cdot),\ e) \\ G_2(x(\cdot),\ (\psi\cdot),\ u(\cdot),\ e) \\ G_3(x(\cdot),\ \psi(\cdot),\ u(\cdot),\ e) \\ g(x(t_1)) \end{bmatrix},$$

显然 $G:B\to B$. 由此，求解(16)—(19)就等价于在 Banach 空间 B 内求解函数方程

$$G(z)=0.$$

易见 $G(\cdot)$ 是 B 内的连续可微函数. 根据前面讲述过的 Newton 法，求解此方程转化为求近似解序列 $\{z_i\}$，而为了求出此序列，首先应求出 $\Delta G(z_i)$ 关于 δz_i 的线性主部 $\delta G(z_i)=(\delta G_1(z_i),\delta G_2(z_i),\ \delta G_3(z_i),\ \delta g(z_i))^T$，然后建立相应于式(11)的下述迭代方程组：

$$\left.\begin{aligned} &G_1(z_i)+\delta G_1(z_i)=0, \\ &G_2(z_i)+\delta G_2(z_i)=0, \\ &G_3(z_i)+\delta G_3(z_i)=0, \\ &g(z_i)+\delta g(z_i)=0. \end{aligned}\right\} \qquad (20)$$

现在计算 $\delta G(z_i)$. 从式(7)易见有

$$\delta G(z)=\lim_{\varepsilon\to 0}\frac{G(z+\varepsilon\delta z)-G(z)}{\varepsilon}.$$

注意到 $\qquad z_i=(x_i(\cdot),\ \psi_i(\cdot),\ u_i(\cdot),\ e_i)^T,$

$$\delta z_i=z_{i+1}-z_i=(\delta x_i(\cdot),\ \delta\psi_i(\cdot),\ \delta u_i(\cdot),\ \delta e_i)^T,\ t_0\leqslant t\leqslant t_1.$$

于是

$$\delta G_1(z_i)=\left[\frac{\partial}{\partial\varepsilon}G_1(z_i+\varepsilon\delta z_i)\right]_{\varepsilon=0}$$

$$=\int_{t_0}^{t}\left\{\left(\frac{\partial f(\tau,\ x_i(\tau),\ u_i(\tau))}{\partial x}\right)^T\delta x_i(\tau)\right.$$

$$+\left(\frac{\partial f(\tau,\ x_i(\tau),\ u_i(\tau))}{\partial u}\right)^T \delta u_i(\tau)\Big\}d\tau-\delta x_i(t), \tag{21}$$

$$\begin{aligned}\delta G_2(z_i)&=\left[\frac{\partial}{\partial \varepsilon}G_2(z_i+\varepsilon\delta z_i)\right]_{\varepsilon=0}\\&=\frac{\partial g(x_i(t_1))}{\partial x}\delta e_i+\left(\frac{\partial}{\partial x}\left[\frac{\partial g(x_i(t_1))}{\partial x}e_i\right]\right)^T\delta x_i(t_1)\\&\quad+\int_{t_1}^{t}\Big\{-\frac{\partial f(\tau,\ x(\tau),\ u(\tau))}{\partial x}\delta\psi_i(\tau)\\&\quad-\left(\frac{\partial}{\partial x}\left[\frac{\partial f(\tau,\ x_i(\tau),\ u_i(\tau))}{\partial x}\psi_i(\tau)\right]\right)^T\delta x_i(\tau)\\&\quad-\left(\frac{\partial}{\partial u}\left[\frac{\partial f(\tau,\ x_i(\tau),\ u_i(\tau))}{\partial x}\psi_i(\tau)\right]\right)^T\delta u_i(\tau)\\&\quad+\left(\frac{\partial^2}{\partial x^2}f^0(\tau,\ x_i(\tau),\ u_i(\tau))\right)^T\delta x_i(\tau)\\&\quad+\left(\frac{\partial^2}{\partial u\partial x}f^0(\tau,\ x_i(\tau),\ u_i(\tau))\right)^T\delta u_i(\tau)\Big\}d\tau\\&\quad-\delta\psi_i(t),\end{aligned} \tag{22}$$

$$\begin{aligned}\delta G_3(z_i)&=\left[\frac{\partial}{\partial \varepsilon}G_3(z_i+\varepsilon\delta z_i)\right]_{\varepsilon=0}\\&=-\left(\frac{\partial^2 f^0(t,\ x_i(t),\ u_i(t))}{\partial x\partial u}\right)^T\delta x_i(t)\\&\quad-\left(\frac{\partial^2 f^0(t,\ x_i(t),\ u_i(t))}{\partial u^2}\right)^T\delta u_i(t)\\&\quad+\left(\frac{\partial}{\partial u}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t)\right]\right)^T\delta x_i(t)\\&\quad+\left(\frac{\partial}{\partial u}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t)\right]\right)^T\delta u_i(t)\\&\quad+\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\delta\psi_i(t),\end{aligned} \tag{23}$$

$$\begin{aligned}\delta g(x_i(t_1))&=\left[\frac{\partial}{\partial \varepsilon}g(x_i(t_1)+\varepsilon\delta x_i(t_1))\right]_{\varepsilon=0}\\&=\left(\frac{\partial g(x_i(t_1))}{\partial x}\right)^T\delta x_i(t).\end{aligned} \tag{24}$$

上面诸式中 $\delta x_i(t)=x_{i+1}(t)-x_i(t)$, $\delta\psi_i(t)=\psi_{i+1}(t)-\psi_i(t)$, $\delta u_i(t)=u_{i+1}(t)-u_i(t)$, $\delta e_i=e_{i+1}-e_i$. 将(21)—(24)代入(20),就得到关于求解最优控制问题(12)—(14)的近似解序列 $\{z_i\}$, $z_i=(x_i(\cdot),\ \psi_i(\cdot),\ u_i(\cdot),\ e_i)^T$,的迭代方程组了.

下面将进一步把(20)简化成关于 $\delta x_i(t)$ 和 $\delta\psi_i(t)$ 的迭代形式的微分方程组边值问题. 为此, 对方程组(20)两边求导, 注意到 $G(z_i)$ 和 $\delta G(z_i)$ 的上述诸具体表达式,整理之即得

$$
\begin{aligned}
\frac{d}{dt}\delta x_i(t)=&\left(\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial x}\right)^T\delta x_i(t)\\
&+\left(\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\right)^T\delta u_i(t)\\
&+f(t,\ x_i(t),\ u_i(t))-\frac{dx_i(t)}{dt},
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
\frac{d}{dt}\delta\psi_i(t)=&\left\{\frac{\partial^2}{\partial x^2}f^0(t,\ x_i(t),\ u_i(t))\right.\\
&\left.-\frac{\partial}{\partial x}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial x}\psi_i(t)\right]\right\}^T\delta x_i(t)\\
&-\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial x}\delta\psi_i(t)\\
&+\left\{\frac{\partial^2}{\partial u\partial x}f^0(t,\ x_i(t),\ u_i(t))\right.\\
&\left.-\frac{\partial}{\partial u}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial x}\psi_i(t)\right]\right\}^T\delta u_i(t)\\
&+\frac{\partial f^0(t,\ x_i(t),\ u_i(t))}{\partial x}\\
&-\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial x}\psi_i(t)-\frac{d}{dt}\psi_i(t),
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
&\left\{-\frac{\partial^2}{\partial x\partial u}f^0(t,\ x_i(t),\ u_i(t))\right.\\
&\quad\left.-\frac{\partial}{\partial x}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t)\right]\right\}^T\delta x_i(t)\\
&\quad+\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\delta\psi_i(t)+\left\{-\frac{\partial^2}{\partial u^2}f^0(t,\ x_i(t),\ u_i(t))\right.
\end{aligned}
$$

$$+\left(\frac{\partial}{\partial u}\left[\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t)\right]\right)\Big\}^T\delta u_i(t)$$

$$=\frac{\partial}{\partial u}f^0(t,x_i(t),u_i(t))-\frac{\partial f(t,\ x_i(t),\ u_i(t))}{\partial u}\psi_i(t). \tag{27}$$

式(27)表明 $\delta x_i(t)$、$\delta\psi_i(t)$ 和 $\delta u_i(t)$ 之间是线性关系．假设从(27)可唯一地解出

$$\delta u_i(t)=L_i(t)\delta x_i(t)+M_i(t)\delta\psi_i(t)+N_i(t), \tag{28}$$

此处 $L_i(t)$、$M_i(t)$ 和 $N_i(t)$ 均只是与 $x_i(t)$、$\psi_i(t)$ 和 $u_i(t)$ 有关的确定的函数，则把(28)代入(25)和(26)，整理后可见

$$\frac{d}{dt}\delta x_i(t)=A_i(t)\delta x_i(t)+B_i(t)\delta\psi_i(t)+E_i(t), \tag{29}$$

$$\frac{d}{dt}\delta\psi_i(t)=C_i(t)\delta x_i(t)+D_i(t)\delta\psi_i(t)+F_i(t), \tag{30}$$

其中 A_i、B_i、C_i、D_i、E_i 和 F_i 均系 $x_i(t)$、$\psi_i(t)$ 和 $u_i(t)$ 的已知函数．

注意到 $G(z_i)$ 和 $\delta G(z_i)$ 的具体表示式，从方程组(20)可见 $\delta x_i(t)$，$\delta\psi_i(t)$ 的边值条件是

$$\delta x_i(t_0)=\bar{x}_{i0},\qquad G_i^T(t_1)\delta x_i(t_1)=-g_i, \tag{31}$$

$$\delta\psi_i(t_1)=G_i(t_1)\delta e_i+H_i(t_1)\delta x_i(t_1)+h_i, \tag{32}$$

此处

$$G_i(t_1)=\frac{\partial}{\partial x}g(x_i(t_1)),$$

$$H_i(t_1)=\left(\frac{\partial}{\partial x}\left[\frac{\partial g(x_i(t_1))}{\partial x}e_i\right]\right)^T,$$

$$\bar{x}_{i0}=x_0-x_i(t_0),\qquad g_i=g(x_i(t_1)),\qquad h_i=G_i(t_1)e_i.$$

式(28)—(32)是关于 $\delta x_i(t)$、$\delta\psi_i(t)$、$\delta u_i(t)$ 和 δe_i 的迭代微分方程边值问题．在已求得 $\delta x_{i-1}(t)$、$\delta\psi_{i-1}(t)$、$\delta u_{i-1}(t)$ 和 δe_{i-1} 后，就可得到 $x_i(t)$、$\psi_i(t)$、$u_i(t)$ 和 e_i; 于是方程组(28)—(32)中的各个系数均为已知，然后求解微分方程边值问题(29)—(32)而得 $\delta x_i(t)$、$\delta\psi_i(t)$ 和 δe_i，再从(28)就得 $\delta u_i(t)$．这样，即可求出

$x_{i+1}(t)=x_i(t)+\delta x_i(t)$, $\psi_{i+1}(t)=\psi_i(t)+\delta\psi_i(t)$, $u_{i+1}(t)=u_i(t)+\delta u_i(t)$, $e_{i+1}=e_i+\delta e_i$; 重复迭代，就求出满足最优控制问题必要条件 (15)～(18) 的近似解序列 $\{x_i(t), \psi_i(t), u_i(t), e_i\}$, $i=0, 1, 2, \cdots$; 从而有可能求得最优控制问题(12)～(14)的解.

总之，求解最优控制问题(12)～(14)转化为求解形式如下的微分方程边值问题：

$$\frac{d}{dt}x(t)=A(t)x(t)+B(t)\psi(t)+E(t), \tag{33}$$

$$\frac{d}{dt}\psi(t)=C(t)x(t)+D(t)\psi(t)+F(t), \tag{34}$$

$$x(t_0)=\bar{x}_0, \tag{35}$$

$$G^T x(t_1)=-g, \tag{36}$$

$$\psi(t_1)=Ge+Hx(t_1)+h, \tag{37}$$

这里 G 是 $s\times n$ 满秩常阵，H 是 $s\times n$ 常阵，g 和 h 分别是 s 和 n 维常向量. 因此，下面专就求解(33)～(37)的微分方程边值问题进行研究.

解线性微分方程边值问题是一个困难问题，并非都能求解，下面仅在某个 Riccati 方程(见式(41))的解可具体求得的假设下来解此问题.

设微分方程组(33)和(34)的解 $x(t)$, $\psi(t)$ 满足线性关系：

$$\psi(t)=k(t)x(t)+l(t), \quad t\in[t_0, t_1], \tag{38}$$

这里 $k(t)$ 是 $n\times n$ 矩阵，$l(t)\in\mathbb{R}^n$，则易推出必有关系式

$$\left[\frac{dk(t)}{dt}+k(t)A(t)+k(t)B(t)k(t)-D(t)k(t)-C(t)\right]x(t)$$
$$+\frac{dl(t)}{dt}+[k(t)B(t)-D(t)]l(t)+k(t)E(t)-F(t)=0 \tag{39}$$

成立. 反之，若取 $k(t)$ 和 $l(t)$, $t\in[t_0, t_1]$，满足微分方程组：

$$\frac{d}{dt}k(t) = -k(t)A(t) - k(t)B(t)k(t) + D(t)k(t) + C(t), \tag{40}$$

$$\frac{d}{dt}\ l(t) = [-k(t)B(t) + D(t)]l(t) - k(t)E(t) + F(t), \tag{41}$$

则式(39)必成立；进一步作线性变换

$$\psi = k(t)x + l(t),$$

代入(33)，求微分方程

$$\frac{dx}{dt} = A(t)x + B(t)[k(t)x + l(t)] + E(t),$$

的解 $x(t)$，则可直接验证

$$\psi(t) = k(t)x(t) + l(t)$$

是方程(34)的解，即这样求得的 $x(t)$，$\psi(t)$ 必定是微分方程组(33)和(34)的解．若还取 $k(t)$ 和 $l(t)$ 的边值为

$$k(t_1) = H,\ l(t_1) = Ge + h, \tag{42}$$

或为

$$k(t_1) = 0, \quad l(t_1) = Ge + Hx(t_1) + h, \tag{43}$$

则显然边值条件(37)成立．

这样，若 Riccati 矩阵微分方程(40)在终值条件(42)或(43)之下有解的话，则求解微分方程边值问题(33)～(37)简化为：首先在终值条件(42)或(43)之下求解(40)和(41)；其次，代(38)到(33)按边界条件求解 $x(t)$；最后把 $x(t)$ 代入(38)就得到 $\psi(t)$．现按此步骤来进行计算．

我们首先求解(40)、(41)，只就边值条件(42)讨论．在终值条件 $k(t_1) = H$ 之下，求解出 Riccati 微分方程(40)的解 $k(t)$，$t \in [t_0, t_1]$；代入(41)，求此方程所对应的矩阵齐次微分方程

$$\begin{cases} \dfrac{d}{dt}\ L(t) = [-k(t)B(t) + D(t)]\,L(t), \\ L(t_1) = I \end{cases}$$

2 九四七六

的解 $L(t)$，其中 I 是 $n\times n$ 单位矩阵. 于是就得到方程(41)的满足边界条件(42)的特解

$$l(t)=L(t)\left\{(Ge+h)+\int_{t_1}^{t}L^{-1}(s)[F(s)-k(s)E(s)]ds\right\}$$
$$=L(t)(Ge+h)+l_0(t),\quad t\in[t_0,\ t_1],\tag{44}$$

此处

$$l_0(t)\triangleq\int_{t_1}^{t}L(t)L^{-1}(s)[F(s)-k(s)E(s)]ds.$$

其次，在初值(35)之下求(33)的解 $x(t)$, $t\in[t_0,\ t_1]$. 为此，用(38)代入(33)得

$$\frac{d}{dt}x(t)=[A(t)+B(t)k(t)]x(t)+B(t)l(t)+E(t).\tag{45}$$

求此方程所对应的矩阵齐次方程

$$\begin{cases}\dfrac{d}{dt}X(t)=[A(t)+B(t)k(t)]X(t),\\ X(t_0)=I\end{cases}$$

的解 $X(t)$, $t\in[t_0,\ t_1]$，其中 I 是$n\times n$ 阶单位矩阵. 于是(45)的满足初值(35)的解

$$x(t)=X(t)\left\{\bar{x}_0+\int_{t_0}^{t}X^{-1}(s)[B(s)l(s)+E(s)]ds\right\}.\tag{46}$$

用式(44)的 $l(t)$ 代入，得

$$x(t)=X(t)\bar{x}_0+\left(\int_{t_0}^{t}X(t)X^{-1}(s)B(s)L(s)ds\right)Ge$$
$$+\int_{t_0}^{t}X(t)X^{-1}(s)[B(s)(L(s)h+l_0(s))+E(s)]ds,\tag{47}$$

$x(t)$ 的表达式中还含有待定的常向量 e，这可利用边值条件 (36)

来确定。

为此，在上式中令 $t=t_1$，且引入记号

$$\begin{cases} x_1 \triangleq X(t_1)\bar{x}_0, \\ w_1 \triangleq \int_{t_0}^{t_1} X(t_1)X^{-1}(s)B(s)L(s)ds, \\ v_1 \triangleq \int_{t_0}^{t_1} X(t_1)X^{-1}(s)[B(s)(L(s)h+l_0(s))+E(s)]ds, \end{cases} \tag{48}$$

则
$$x(t_1)=x_1+v_1+w_1Ge$$

应用边值条件(36)就得

$$-g=G^Tx(t_1)=G^Tx_1+G^Tv_1+G^Tw_1Ge,$$

从而

$$e=-(G^Tw_1G)^{-1}(g+G^Tx_1+G^Tv_1). \tag{49}$$

最后，把(49)的 e 代入(47)和(44)，就得 $x(t)$，$l(t)$；把 $x(t)$，$l(t)$ 以及 $k(t)$ 代入(38)就得 $\psi(t)$，从而求得边值问题(33)—(37)的解 $(x(t), \psi(t))$。

综合上述，我们得到了求解边值问题(33)—(37)如下的：

算法 2

第 1 步 求 Riccati 方程(40)具终值 $k(t_1)=\Pi$ 的解 $k(t)$，$t\in[0, T]$。

第 2 步 求具终值条件(42)的非齐次线性微分方程(41)的解(44)：

$$l(t)=L(t)(Ge+h)+l_0(t), \quad t\in[t_0, t_1].$$

第 3 步 用公式(38)的 $\psi(t)$ 代入(33)而求此关于 $x(t)$ 的非齐次线性微分方程的在初值条件(35)之下的解（按公式(47)）。

$$x(t)=X(t)\bar{x}_0+\left\{\int_{t_0}^{t} X(t)X^{-1}(s)B(s)L(s)ds\right\}Ge$$
$$+\int_{t_0}^{t} X(t)X^{-1}(s)\{B(s)(L(s)h+l_0(s))+E(s)\}ds.$$

第 4 步　按公式(49)，即按

$$e=-(G^T w_1 G)^{-1}(g+G^T x_1+G^T v_1)$$

求出参数向量 e，此处 x_1，v_1 和 w_1 按式(48)求出。

第 5 步　按公式(38)求出 $\psi(t)$，$t\in[t_0, t_1]$。

由上述可见，算法 2 的关键在于是否能求出 Riccati 方程(40)的满足终值条件(42)或(43)的解．下面，我们将指出，对线性二次最优控制问题这种重要的特殊情形，相应于方程(40)的 Riccati 方程的解是能求出来的。

现在考虑系统的状态方程是

$$\frac{d}{dt}x(t)=Ax(t)+Bu(t),\quad t\in[0, T], \tag{50}$$

$$x(t_0)=x_0,$$

性能指标为

$$J(u(\cdot))=\frac{1}{2}\langle Q_1x(T), x(T)\rangle+\frac{1}{2}\int_0^T(\langle Qx(t), x(t)\rangle+\langle Ru(t), u(t)\rangle)dt, \tag{51}$$

此处 $x\in\mathbb{R}^n$，$u(\cdot)\in L^2([0, T], \mathbb{R}^m)$，$x_0\in\mathbb{R}^n$ 给定；A 和 B 分别是 $n\times n$ 和 $n\times m$ 常阵，Q_1 和 Q 是 $n\times n$ 对称半正定常阵，R 是 $m\times m$ 对称正定常阵。

设 $u(\cdot)$ 和 $x(\cdot)$ 是最优解，$\psi(\cdot)$ 是最大值原理中对应的伴随方程的解，则从最大值原理或直接从定理 III-4-2 知，$u(t)$、$x(t)$ 和 $\psi(t)$ 必满足关系：

$$\frac{d}{dt}x(t)=Ax(t)+Bu(t), \tag{52}$$

$$\frac{d}{dt}\psi(t)=Qx(t)-A^T\psi(t), \tag{53}$$

$$x(0)=x_0,\quad \psi(T)=-Q_1x(T), \tag{54}$$

$$-Ru(t)+B^T\psi(t)=0, \tag{55}$$

由此知 $x(\cdot)$, $\psi(\cdot)$ 必为定常线性微分方程组

$$\frac{d}{dt}x(t)=Ax(t)+BR^{-1}B^T\psi(t), \tag{56}$$

$$\frac{d}{dt}\psi(t)=Qx(t)-A^T\psi(t), \tag{57}$$

$$x(0)=x_0, \quad \psi(T)=-Q_1x(T) \tag{58}$$

的解．也就是说，求解线性系统二次最优控制问题(50)，(51)就转化为求解(56)—(58)的边值问题．

从上面的算法或直接应用定理 III-4-3 可见，相应于边值问题(56)、(57)和(58)的 Riccati 方程是

$$\frac{d}{dt}k(t)=-k(t)A-k(t)BR^{-1}B^Tk(t)-A^Tk(t)+Q, \tag{59}$$

$$P(T)=-Q_1. \tag{60}$$

下面我们将指出，完全能求出此 Riccati 方程的解，从而按算法2，就可求出边值问题(56)、(57)和(58)的解 $x(t)$, $\psi(t)$，即线性二次型最优控制问题(50)、(51)的解了．

注意到式(56)和(57)是线性齐次微分方程组，为求满足边值(58)的特解，首先，求(56)和(57)的标准基本解矩阵 $Y(t,\tau)$, $t\in[0,T]$, $Y(\tau,\tau)=I$ 是 $2n\times 2n$ 单位矩阵．记

$$Y(t,\tau)=\begin{pmatrix}Y_{11}(t,\tau) & Y_{12}(t,\tau)\\ Y_{21}(t,\tau) & Y_{22}(t,\tau)\end{pmatrix}$$

此处 Y_{ij} 是 $n\times n$ 子矩阵，则微分方程组(56)和(57)的任定的一个解 $(x(t),\psi(t))$ 为

$$\begin{pmatrix}x(t)\\ \psi(t)\end{pmatrix}=Y(t,\tau)\begin{pmatrix}x(\tau)\\ \psi(\tau)\end{pmatrix},$$

其中 $t=\tau$ 时的初值为 $(x(\tau),\psi(\tau))$，即有

$$\begin{cases} x(t)=Y_{11}(t,\ \tau)x(\tau)+Y_{12}(t,\ \tau)\psi(\tau), \\ \psi(t)=Y_{21}(t,\ \tau)x(\tau)+Y_{22}(t,\ \tau)\psi(\tau), \quad t、\tau\in[0,T]. \end{cases} \tag{61}$$

设$(x(t),\ \psi(t))$是满足边值条件(58)的第2式的特解，则

$$\begin{aligned} 0&=\psi(T)+Q_1x(T) \\ &=[Q_1Y_{11}(T,\ \tau)+Y_{21}(T,\ \tau)]x(\tau) \\ &\quad+[Q_1Y_{12}(T,\ \tau)+Y_{22}(T,\ \tau)]\psi(\tau). \end{aligned}$$

改τ为t，就得

$$K(t)\psi(t)=-[Q_1Y_{11}(T,\ t)+Y_{21}(T,\ t)]x(t), \quad t\in[0,\ T], \tag{62}$$

此处

$$K(t)\triangleq Q_1Y_{12}(T,\ t)+Y_{22}(T,\ t), \quad t\in[0,\ T]. \tag{63}$$

显然，若$K(t)$，$t\in[0,\ T]$，非奇异，则式(62)可写为下面的形式

$$\psi(t)=k(t)x(t), \quad t\in[0,\ T], \tag{64}$$

此处

$$k(t)\triangleq -K^{-1}(t)[Q_1Y_{11}(T,\ t)+Y_{21}(T,\ t)], \quad t\in[0,\ T]. \tag{65}$$

即如果$K(t)$，$t\in[0,\ T]$，非奇异，则微分方程组(56)和(57)的满足边值(58)的第二式的解$(x(t),\ \psi(t))$中，$\psi(t)$必为$x(t)$的线性齐次形式。又易推算出：若$(x(t),\ \psi(t))$是微分方程组(56)和(57)的解，且$\psi(t)$是$x(t)$的齐次线性形(64)的话，则$k(t)$必为Riccati方程(59)的解。注意到$Y(T,\ T)=I$，对式(65)的$k(t)$可检验得知

$$k(T)=-Q_1.$$

这样，即得

定理1 若$K(t)$，$t\in[0,\ T]$，非奇异，则Riccati方程(59)必有具初始条件(60)的解(65)。

下面证明：当$t\in[0,T]$时，$K(t)$确是非奇异的。显然，$K(T)=$

$Y_{22}(T, T)$ 为 $n\times n$ 单位矩阵，故为非奇异的．下面只要指出，当 $t\in[0, T)$ 时，$K(t)$ 为非奇异即可．采用矛盾法，设存在某时刻 $t'\in[0, T)$ 使 $\det K(t')=0$，则齐次代数方程组

$$K(t')\psi=0$$

存在非零解 ψ'，取线性齐次微分方程组(56)和(57)的在 t' 时刻的初值为 $(0, \psi')$ 的特解 $(\bar{x}(t), \bar{\psi}(t))$，$t\in[0, T]$，它必定不是平凡解，且从(61)知

$$\begin{aligned}\bar{x}(t)&=Y_{12}(t, t')\psi',\\ \bar{\psi}(t)&=Y_{22}(t, t')\psi'',\end{aligned}\qquad t\in[0, T],$$

故可见

$$\begin{aligned}Q_1\bar{x}(T)+\bar{\psi}(T)&=Q_1Y_{12}(T, t')\psi'+Y_{22}(T, t')\psi'\\ &=K(t')\psi'=0.\end{aligned}\tag{66}$$

另一方面，注意到微分方程组(56)，(57)右端主对角线子矩阵的特性，可算得：

$$\frac{d}{dt}\langle\bar{x}(t), \bar{\psi}(t)\rangle=\langle\bar{x}(t), Q\bar{x}(t)\rangle+\langle\bar{\psi}(t), BR^{-1}B^T\bar{\psi}(t)\rangle,$$

故

$$\begin{aligned}\langle\bar{x}(T), \bar{\psi}(T)\rangle=\int_{t'}^{T}&\langle\bar{x}(t), Q\bar{x}(t)\rangle\\ &+\langle\bar{\psi}(t), BR^{-1}B^T\bar{\psi}(t)\rangle dt.\end{aligned}\tag{67}$$

用式(66)的 $\bar{\psi}(T)=-Q_1\bar{x}(T)$ 代入上式，并注意到 Q_1, Q 和 $BR^{-1}B^T$ 均为正定或半正定，即知：恒等式(67)的成立必得

$$\langle\bar{x}(T), Q_1\bar{x}(T)\rangle=\langle\bar{x}(t), Q\bar{x}(t)\rangle=\langle\bar{\psi}(t), BR^{-1}B^T\bar{\psi}(t)\rangle=0,$$
$$t\in[0, T],$$

从而必有

$$Q_1\bar{x}(T)=Q\bar{x}(t)=BR^{-1}B^T\bar{\psi}(t)=0,\quad t\in[0, T].$$

这样，$(\bar{x}(t), \bar{\psi}(t))$ 所满足的微分方程组(56)和(57)取形式

$$\frac{d\bar{x}(t)}{dt}=A\bar{x}(t),$$
$$\frac{d\bar{\psi}(t)}{dt}=-A^T\bar{\psi}(t).$$

从而

$$\begin{aligned}&\bar{x}(t)=\bar{x}(t')\exp[At]=0,\\&\bar{\psi}(t)=\bar{\psi}(t')\exp\{-A^Tt\}=\psi'\exp[-A^Tt],\qquad t\in[0,T].\end{aligned}\tag{68}$$

注意到 $\psi'\neq 0$，故得 $\bar{\psi}(T)\neq 0$；但从式(68)及(66)却见 $\bar{\psi}(T)=0$，所得矛盾证明了 $K(t)\neq 0$，$t\in[0,T]$，从而有下面的

定理 2 由式(63)所得的 $K(t)$，$t\in[0,T]$，是非奇异的.

综上所述，线性二次最优控制问题相对应的 Riccati 方程的解是可求得的，从而算法 2 可以实现.

§4 解约束最优控制问题的罚函数方法

实际问题中，控制和状态通常是受约束的．下面我们考虑微分方程组

$$\begin{cases}\frac{d}{dt}x(t)=f(t,x(t),u(t)),\\x(t_0)=x_0,\end{cases}\tag{1}$$

和性能指标

$$J(u(\cdot))=\int_{t_0}^{t_1}f^0(t,x(t),u(t))dt,\tag{2}$$

这里 $f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$，$f^0:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 是连续的，且 $\frac{\partial f^0}{\partial x}$，$\frac{\partial f^0}{\partial u}$，$\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial u}$ 也连续，t_0，t_1 和 x_0 认为是给定的.

最优控制问题 取 $u(\cdot)\in L^\infty\triangleq L^\infty([t_0,t_1],\mathbb{R}^m)$ 的有界可测函数的等价类空间，使得微分方程组(1)的解 $x(t)=x(t,u(\cdot))$，$x(t_0)=x_0$，满足下列一种或多种同时具有的约束，并使泛函(2)取

最小值. 这些约束是, 当 $t\in[t_0, t_1]$ 时,

$$g(x(t_1))=0, \tag{3}$$

$$q(t, u(t))\leqslant 0, \tag{4}$$

$$r(t, x(t))=0, \tag{5}$$

$$s(t, x(t))\leqslant 0, \tag{6}$$

$$w(t, x(t), u(t))\leqslant 0, \tag{7}$$

这里 $g:\mathbb{R}^n\to\mathbb{R}^\alpha$, $q:\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}^\beta$, $r:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^\gamma$, $s:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^\mu$ 和 $w:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^\nu$ 均为其变元的连续函数.

还可提出对控制 $u(\cdot)$ 和状态 $x(\cdot)$ 的别种形式的约束. 上述的这个约束最优控制问题可以化为如下形式的非线性规划最优化问题: 求

$$\inf\{J(u)\,|\,u(\cdot)\in L^\infty, \text{且有下列一种或多种约束:}$$
$$g(x(t_1, u))=0,\ q(t, u(t))\leqslant 0,\ r(t, x(t, u))=0,$$
$$s(t, x(t, u))\leqslant 0,\ w(t, x(t, u), u(t))\leqslant 0, \quad \text{a.e.}\ t\in[t_0, t_1]\}, \tag{8}$$

此处 $x(t)=x(t, u)$ 是微分方程(1)的解, $J(u)$ 是性能指标(2).

对于控制和状态均不受限制时的最优控制问题, 在 § 2 和 § 3 中, 我们曾把它转化为等价的最优化问题:

$$\min\{J(u(\cdot))\,|\,u(\cdot)\in L^\infty\},$$

然后用梯度方法或牛顿法求其解. 此处

$$J(u(\cdot))=\int_{t_0}^{t_1} f^0(t, x(t, u(\cdot)), u(t))dt,$$

$x(t, u(\cdot))$ 是微分方程(1)的解, 对于上述约束最优控制问题, 就不能直接再用 § 2, § 3 的方法了, 但人们自然地考虑, 是否能把问题转化为求解无约束的最优控制问题呢? 早在四十年代, R. Courant 就曾在普通微积分学中, 为求解满足边值条件 $\Psi(p)=0$($\Psi(p)\geqslant 0$) 的函数 $\Phi(p)$ 的极小值问题 A, 转而解无约束的函数 $\Phi_k(p)=$

$\Phi(p)+k\Psi(p)$的极小值问题 A_k，在一定条件下，问题 A_k 的解，即极小值点 p_k，所构成的序列的极限点 p_0 必为原问题 A 的极小值点．这样，把求解约束最优化问题转化为求解无约束最优化问题序列及其收敛性的问题了，这是微积分学中的罚函数方法．我们下面，将这种罚函数方法推广到无限维空间 $L^\infty([t_0, t_1], \mathbb{R}^m)$内求解约束最优化问题，然后应用它去求解(1)～(7)的约束最优控制问题．

上面我们已指出最优控制问题(1)～(7)转化为形如(8)的约束最优化问题了，现在进一步把问题的提法精确化．对于(3)～(7)的各个约束，记

$$B_1=\{u(\cdot)\,|\,u(\cdot)\in L^\infty,\ g(x(t_1,\ u(\cdot)))=0,\ \},\tag{9}$$

$$B_2=\{u(\cdot)\,|\,u(\cdot)\in L^\infty,\ q(t,\ u(t))\leqslant 0,\ \text{a.e.}t\in[t_0,\ t_1]\}\tag{10}$$

$$B_3=\{u(\cdot)\,|\,u(\cdot)\in L^\infty,\ r(t,\ x(t,\ u(\cdot)))=0,\ \text{a. e.}\ t\in[t_0,\ t_1]\},\tag{11}$$

$$B_4=\{u(\cdot)\,|\,u(\cdot)\in L^\infty,\ s(t,\ x(t,\ u(\cdot)))\leqslant 0,\ \text{a.e.}t\in[t_0,\ t_1]\},\tag{12}$$

$$B_5=\{u(\cdot)\,|\,u(\cdot)\in L^\infty,\ w(t,\ x(t,\ u(\cdot)),\ u(t))\leqslant 0,\ \text{a.e.}t\in[t_0,\ t_1]\}\tag{13}$$

可验证得 B_i, $i=1$, …, 5 均为 Banach 空间 $L^\infty\triangleq L^\infty([t_0,\ t_1],\ \mathbb{R}^m)$ 内的闭集．这样，对满足(3)—(7)中的一种或多种约束的最优控制问题可转化为研究形式如下的最优化问题：求

$$(\text{CP})\qquad \inf\{J(u)\,|\,u\in L^\infty,\ u\in B\},\tag{14}$$

此处 $J(u)$ 是映 L^∞到 $\mathbb{R}$ 的连续映象，B 是 L^∞ 内某闭集．

下面我们介绍求解 Banach 空间内约束最优化问题(14)的罚

函数方法。

记

$$J_* = \inf\{J(u) \mid u \in B\}, \quad J_0 = \inf\{J(u) \mid u \in L^\infty\},$$

若 B 具非空内部 $\mathring{B}$ 时，还记

$$\mathring{J} = \inf\{J(u) \mid u \in \mathring{B}\}.$$

显然，$J_0 \leqslant J_* \leqslant \mathring{J}$. 右边不等式只在 $\mathring{B}$ 非空时出现. 以后要指出 J_0 是有限数.

若点 $u_* \in B$ 使 $J(u_*) = J_*$ 时，称 u_* 为(CP)的解；若 $\mathring{B}$ 非空，$u_* \in \mathring{B}$ 时，称 u_* 为(CP)的一个非受限解.

下面我们将统一地讲述内罚函数和外罚函数方法，约定在本节里从下面起，在叙述中出现大方括号[ ]时，括号内的和紧靠括号前方的符号、文字分别对应于外罚函数和内罚函数，其余的则两者共同使用(除非明确指出只对其中的一种罚函数使用).

定义 1 设 $\mathring{B}[B]$ 非空，凡具有下列三个性质的函数列 $\{p_i(u)\}$ 称为关于集 B 的内[外]罚函数列：

1° 在 $\mathring{B}[L^\infty]$ 内连续非负；

2° 对 $\mathring{B}[L^\infty]$ 内任一点 u 均有

$$\overline{\lim_{i \to +\infty}}\, p_i(u) = 0;$$

3° 对任给定的点列 $\{u_j\} \subset \mathring{B}[L^\infty \setminus B]$，$u_j \to u_* \in \partial B[L^\infty \setminus B]$，均使下式成立：

$$\lim_{j \to +\infty} p_i(u_j) \ \left[\lim_{\substack{i \to +\infty \\ j \to +\infty}} p_i(u_j)\right] = +\infty.$$

现考虑无约束的最优化问题

$$(\text{UPI})_i \qquad \inf\{J_i(u) \mid u \in \mathring{B}\}, \qquad i = 0, 1, 2, \cdots, \tag{15}$$

$$(\text{UPE})_i \qquad \inf\{J_i(u) \mid u \in L^\infty\}, \qquad i = 0, 1, 2, \cdots, \tag{16}$$

此处

$$J_i(u) = J(u) + p_i(u), \qquad i = 0, 1, 2, \cdots. \tag{17}$$

记
$$\mathring{J}_i=\inf\{J_i(u)\mid u\in\mathring{B}\},$$
$$J_i=\inf\{J_i(u)\mid u\in L^\infty\}.$$
在 $\mathring{J}_i$ 的定义中，显然要假定 $\mathring{B}$ 非空. 我们还给出:

假设 $\mathring{H}[H]$　设在 $\mathring{B}[B]$ 内存在两点 u'' 和 u'，使 $J(u'')<J(u')$，且集 $Z\triangleq\{u\mid J(u)\leqslant J(u'),\ u\in L^\infty\}$ 是紧致集.

下面我们首先证明约束最优化问题(CP)和无约束最优化问题 $(\mathrm{UPI})_i$ $[(\mathrm{UPE})_i]$，$i=0, 1, \cdots$，的解存在；然后讨论 $(\mathrm{UPI})_i$ $[(\mathrm{UPE})_i]$，$i=0, 1, \cdots$，的解序列 $\{u_i\}$ 的收敛性问题；最后提供若干充分和必要条件，使带罚函数的无约束最优化问题 $(\mathrm{UPI})_i$ $[(\mathrm{UPE})_i]$，$i=0, 1, 2, \cdots$，的解序列 $\{u_i\}$ 的任一聚点必定是原来的约束最优化问题(CP)的解. 这样，就把求解约束最优化问题转化为求解无约束最优化问题序列，并研究其解序列的收敛性问题，从而原约束最优化问题得到解决，这就是罚函数方法的基本思路. 下面，我们按上述思路来讲述.

易证得

引理 1　设 A 和 B 是 L^∞ 内的两个非空集，则
$$\inf_{u\in A\cup B} J(u)=\min\{\inf_{u\in A} J(u),\ \inf_{u\in B} J(u)\}.$$

引理 2　在假设 $\mathring{H}[H]$ 成立下，(CP)必有解，且其任一解必在集 $B\cap Z$ 内.

证　在假设 $\mathring{H}$ 成立下，假设 H 必成立，故只要讨论假设 H 成立的情形. 因 $B=(B\cap Z)\cup(B\cap CZ)$，$CZ\triangleq L^\infty\backslash Z$，故从引理 1 及假设 H 即得
$$\inf\{J(u)\mid u\in B\}=\inf\{J(u)\mid u\in B\cap Z\},$$
而 $B\cap Z$ 是紧致集的闭子集，故也是紧致集，从而(CP)有解，且解在 $B\cap Z$ 内. ∎

引理 3　若假设 $\mathring{H}[H]$ 成立，$\{p_i(u)\}$ 是关于集 B 的内[外]罚

函数列，则 J_0、J_*、$\mathring{J}$ 和 $\mathring{J}_i[J_i]$ 是有限数，且有关系式：

$$J_*[J_0] \leqslant \varliminf_{i\to+\infty} \mathring{J}_i[J_i] \leqslant \varlimsup_{i\to+\infty} \mathring{J}_i[J_i] \leqslant \mathring{J}[J_*]. \qquad (18)$$

证 注意到 $L^\infty = Z \cup CZ$、引理 1 及假设 $\mathring{H}[H]$ 即得

$$J_0 = \inf\{J(u) \mid u \in L^\infty\}$$
$$= \min(\inf\{J(u) \mid u \in Z\}, \inf\{J(u) \mid u \in CZ\})$$
$$= \inf\{J(u) \mid u \in Z\}.$$

因 Z 紧致及 $J(u)$ 连续，故 J_0 是有限数．又注意到 $u' \in \mathring{B}[B] \subset L^\infty$ 得 $J_0 \leqslant J_* \leqslant \mathring{J}$ 再注意到 $p_i(u)$ 非负，任取 $\bar{u} \in \mathring{B}[B]$ 就得

$$\inf_{u\in B} J(u)\ [\inf_{u\in L^\infty} J(u)] \leqslant \inf_{u\in \mathring{B}} J_i(u)\ [\inf_{u\in L^\infty} J_i(u)]$$
$$\leqslant J(\bar{u}) + p_i(\bar{u}),$$

即
$$J_*[J_0] \leqslant \mathring{J}_i[J_i] \leqslant J(\bar{u}) + p_i(\bar{u}) < +\infty.$$

因此，作上、下限的运算，并注意到罚函数定义中的性质 2°，即得

$$J_*[J_0] \leqslant \varliminf_{i\to+\infty} \mathring{J}_i[J_i] \leqslant \varlimsup_{i\to+\infty} \mathring{J}_i[J_i] \leqslant J(\bar{u}),$$

注意到 $\bar{u}$ 是 $\mathring{B}[B]$ 内任一点，就得证式(18)． ▮

引理 4 若假设 $\mathring{H}[H]$ 成立，$\{p_k(u)\}$ 是关于集 B 的内[外]罚函数列，则存在 $N>0$，当 $k>N$ 时，$(\mathrm{UPI})_k[(\mathrm{UPE})_k]$ 有解；且解序列 $\{u_{k*}\}$ 的任一聚点必属集 B．

证 在集 $\mathring{B}[L^\infty]$ 内取点集

$$C_i \triangleq \{u \mid J(u) + p_i(u) \leqslant J(u'),\ u \in \mathring{B}[L^\infty]\}, \quad i = 0, 1, 2\cdots.$$

从罚函数的性质 2° 知，存在 $N>0$，当 $i \geqslant N$ 时，$p_i(u'') < J(u') - J(u'')$，即 $u'' \in C_i$，故 C_i 非空，当 $i \geqslant N$．

现证 C_i 是紧致集．首先，易见 $C_i \subset Z$，从而 $C_i \subset Z \cap \mathring{B}[L^\infty]$．其次，可证 C_i 是闭集，这只要设 $u_j \in C_i$，$u_j \to u_*$，去证 $u_* \in C_i$ 即可．注意到 $C_i \subset Z$，而 Z 紧致，即知 $u_* \in Z$．由于 $J(u)$ 连续和 $p_i(u)$ 在 $\mathring{B}[L^\infty]$ 内连续，故从 $J(u_j) + p_i(u_j) \leqslant J(u')$ 可见

$$J(u_*) + \varliminf_{j\to+\infty} p_i(u_j)\ [p_i(u_*)] \leqslant J(u'), \quad i \geqslant N. \qquad (19)$$

若 $p_i(u)$ 是外罚函数，则由上式可见 $u_* \in C_i$，从而 C_i 是闭集。若 $p_i(u)$ 是内罚函数，注意 $u_j \in C_i \subset \mathring{B}$，得 $u_* \in B$，故 u_* 或属于 B 的内部 $\mathring{B}$ 或属于 B 的边界 ∂B，两者必居其一。但由罚函数的性质 3^0，不可能有 $u_* \in \partial B$，否则

$$J(u_*) + \lim_{j \to +\infty} p_i(u_j) = +\infty,$$

与式(19)相矛盾，故必定 $u_* \in \mathring{B}$，从而 $u_* \in C_i$，即 C_i 是闭集。总之，C_i，$i \geqslant N$，是紧致集 Z 中的闭子集，故 C_i 也是紧致的。

从 $\mathring{B} = C_i + (\mathring{B} \backslash C_i)$ $[L^\infty = C_i \cup (L^\infty \backslash C_i)]$ 及引理1即可见

$$\mathring{J}_i[J_i] = \inf\{J_i(u) \mid u \in C_i\}, \quad i \geqslant N.$$

由于 C_i 紧致，故 C_i 中存在点 u_{i*}，使

$$J_i(u_{i*}) = \mathring{J}_i[J_i], \quad 当\ i \geqslant N, \tag{20}$$

即 $(\text{UPI})_i[(\text{UPE})_i]$ 在 $C_i \subset \mathring{B}[L^\infty]$ 内有解 u_{i*}，$i \geqslant N$。

现证 $\{u_{i*}\}$ 的所有聚点都属于集 B。从上面已可见 $u_{i*} \in Z$，故解序列 $\{u_{i*}\}$ 必有聚点。设 u_* 是其中的一个聚点，则存在 $\{u_{i*}\}$ 的子序列，仍记为 $\{u_{i*}\}$，收敛于 u_*。兹证 $u_* \in B$。若 $\{p_i(u)\}$ 是内罚函数列，则从 $u_{i*} \in \mathring{B}$ 即得 $u_* \in B$。设 $\{p_i(u)\}$ 是外罚函数列，而 $u_* \bar{\in} B$。从引理3知序列 $\{J_i(u_{i*})\} \equiv \{J_i\}$ 有有限上界 J_*；另方面，从外罚函数的性质

$$\lim_{j \to +\infty} p_j(u_{j*}) = +\infty,$$

故必有
$$\lim_{j \to +\infty} J_j(u_j) = +\infty,$$

所得矛盾证明了 $u_* \in B$。 ∎

下面研究 $(\text{UPI})_i[(\text{UPE})_i]$，$i = 0, 1, 2, \cdots$，的解与(CP)的解的关系。

定义2 设 u_* 是(CP)的解，若存在点列 $\{u_j\}$，$u_j \in \mathring{B}$，有性质 $u_j \to u_*$，则称能从集 $\mathring{B}$ 内趋近(CP)的解。

定义 3 若 $\{u_{i*}\}$ 中的任一收敛子序列，仍记为 $\{u_{i*}\}$，及其极限点 u_*，使有关系式

$$\lim_{i\to+\infty} J_i(u_{i*}) = \lim_{i\to+\infty} J(u_{i*}) = J(u_*) = J_*, \tag{21}$$

其中 u_{i*} 是 $(UPI)_i[(UPE)_i]$ 的解，则称当 $i\to+\infty$ 时 $(UPI)_i[(UPE)_i]$ 等价于 (CP)，或当 $i\to+\infty$ 时，$(UPI)_i[(UPE)_i]$ 趋于问题(CP).

从此定义可见，若 $i\to+\infty$ 时，$(UPI)_i[(UPE)_i]$ 等价于(CP)，则可通过 $(UPI)_i[(UPE)_i]$ 的解序列 $\{u_{i*}\}$，求其聚点而得出(CP)的解，且可通过 $\{J_i(u_{i*})\}$ 或 $\{J(u_{i*})\}$ 的极限而求出 $J(u)$ 的极值.

引理 5 若假设 $\mathring{H}$ 成立，则(CP)有能从 $\mathring{B}$ 内趋近的解的充要条件是 $\mathring{J}=J_*$.

证 **充分性** 因假设 $\mathring{H}$ 成立，故 $\mathring{B}\cap Z$ 非空. 从 $\mathring{B}=(\mathring{B}\cap Z)\cup(\mathring{B}\setminus Z)$ 可见 $\mathring{J}=\inf\{J(u)\mid u\in\mathring{B}\cap Z\}$，从而存在序列$\{u_j\}$，$u_j\in\mathring{B}\cap Z$，使

$$\lim_{j\to+\infty} J(u_j) = \mathring{J}. \tag{22}$$

因 Z 是紧致集，故 $\{u_j\}$ 必存在聚点且均属于集 B，任取其中的一个聚点 u_*，仍记 $\{u_j\}$ 收敛于 u_*，这样，从(22)，$\mathring{J}=J_*$ 及 $J(u)$ 的连续性就得

$$J_* = \mathring{J} = \lim_{j\to+\infty} J(u_j) = J(u_*),$$

即 u_* 是(CP)的解.

必要性 设(CP)的解 u_* 可由 $\{u_j\}\subset\mathring{B}$ 趋近，于是从 $J(u)$ 的连续性即得

$$\mathring{J} \leqslant \lim_{j\to+\infty} J(u_j) = J(u_*) = J_*.$$

但显然应有 $J_*\leqslant\mathring{J}$，从而 $\mathring{J}=J_*$. ∎

引理 6 若闭集 B 有非空内部，且 $\overline{\mathring{B}}=B$，则 $\mathring{J}=J_*$，此处 $\overline{\mathring{B}}$

是 $\mathring{B}$ 的闭包.

证 设 u_* 是(CP)的解,即 $J(u_*)=J_*$, $u_*\in B$,则因 $\tilde{\mathring{B}}=B$,故 u_* 的任意小的邻域内均包含有 $\mathring{B}$ 的点;从而存在点列 $u_j\to u_*$, $u_j\in\mathring{B}$. 于是 $\mathring{J}\leqslant J(u_j)$,从 $J(u)$ 的连续性得 $\mathring{J}\leqslant J(u_*)=J_*$. 但应当有 $\mathring{J}\geqslant J_*$,故 $J_*=\mathring{J}$。 ∎

此引理为我们提供了一个判断 $\mathring{J}=J_*$ 的或(CP)有从 $\mathring{B}$ 内趋近的解的法则.

定理 1 若假设 $\mathring{H}$ 成立,$\{p_i(u)\}$ 是关于 B 的内罚函数列,则当 $i\to+\infty$ 时,$(\mathrm{UPI})_i$ 等价于(CP)的充要条件是:或 1) $\mathring{J}=J_*$, 2)(CP)有从 $\mathring{B}$ 内趋近的解.

证 **充分性** 从引理 5 知条件 1) 和 2) 等价,故只要证明条件 1) 的充分性即可.

从引理 4 知,$(\mathrm{UPI})_i$ 有解 u_{i*},且解序列 $\{u_{i*}\}$ 中的任一个聚点 $u_*\in B$,仍记 $u_{i*}\to u_*$. 于是,从 $J(u)$、$p_i(u)$ 的连续性,$p_i(u)$ 的非负性及 $J_i(u_i)=\mathring{J}_i$,并应用引理 3 即得

$$J_*\leqslant J(u_*)=\lim_{i\to+\infty}J(u_{i*})\leqslant\varliminf_{i\to+\infty}J_i(u_{i*})$$

$$\leqslant\varlimsup_{i\to+\infty}J_i(u_{i*})\leqslant\mathring{J},$$

这样,从上面的不等式立即得知条件 1) 和 2) 是当 $i\to+\infty$ 时 $(\mathrm{UPI})_i$ 等价于(CP)的充分条件.

必要性 设 u_{i*} 是 $(\mathrm{UPI})_i$ 的任一个解,u_* 是 $\{u_{i*}\}$ 中任定的一个聚点,且认为 $u_{i*}\to u_*$,则从假设条件知式(21)成立. 另方面,因 $u_{i*}\in\mathring{B}$ 及 $p_i(u)$ 非负,故

$$\mathring{J}\leqslant J(u_{i*})\leqslant J_i(u_{i*}),$$

再从 $\mathring{J}_i=J_i(u_{i*})$,应用引理 3 就得关系式

$$\lim_{i\to+\infty}J_i(u_{i*})=\lim_{i\to+\infty}J(u_{i*})=\mathring{J},$$

将此式与式(21)合并，就得证 $\hat{J}=J_*$. ▌

定理2 若假设 H 成立，$\{p_i(u)\}$ 是关于 B 的外罚函数列，则当 $i\to+\infty$ 时，$(UPE)_i$ 等价于(CP).

证 从引理4知，$(UPE)_i$ 有解 $u_{i*}\in L^\infty$，且 $\{u_{i*}\}$ 中的任一个聚点 $u_*\in B$，为简便起见，仍认为 $u_{i*}\to u_*$. 于是，从 $J(u)$ 的连续性，$p_i(u)$ 的非负性及 $J_i(u_{i*})=J_i$ 应用引理3就得关系式

$$J_*\leqslant J(u_*)=\lim_{i\to+\infty}J(u_{i*})\leqslant\underline{\lim}_{i\to+\infty}J_i(u_{i*})$$

$$\leqslant\overline{\lim}_{i\to+\infty}J_i(u_{i*})\leqslant J_*.$$

因此，立即知式(21)成立，从而本定理得证. ▌

从以上两定理可见，当 $i\to+\infty$ 时，$(UPI)_i[(UPE)_i]$ 等价于(CP)的问题，对内罚函数列来说是有条件的，而对外罚函数列来说则是无条件的.

现在我们采用上述的罚函数方法来解最优控制问题(1)～(7)或即(CP)，就剩下去具体地构造约束(9)～(13)的罚函数了. 在用罚函数去解除(9)～(13)中的约束时，人们只需要解除那些不易处理的等式约束或不等式约束，而不必去解除全部约束. 因此，我们将逐个地就(9)～(13)中的约束来构造罚函数.

对于约束(10)，即集 B_2，我们可按罚函数的定义而构造出关于集 B_2 的内罚函数或外罚函数. 例如，取泛函

$$p''(u(\cdot))=\int_{t_0}^{t_1}\sum_{j=1}^{\beta}(\max\{q^j(t,u(t)),0\})^\delta dt,\qquad \delta\geqslant1.$$

易见，$u\in B_2$ 时，$p''(u)=0$；$u\overline{\in}B_2$ 时，$p''(u)>0$，故取

$$p''_i(u)=\frac{1}{\varepsilon_i}p''(u),\qquad i=0,1,2,\cdots,$$

$\varepsilon_i\to0$，则可见 $\{p''_i(u)\}$ 是关于集 B_2 的外罚数列. 又例如，对集 B_2 设 $q(t,u)=(q^1(t,u),\cdots,q^\beta(t,u))$，$u\in\mathbb{R}^m$，的分量下方有界，即存在常数 $M>0$，使 $-M\leqslant q^j(t,u)$，$j=1,\cdots,\beta$，作泛函

$$p'(u(\cdot))=-\int_{t_0}^{t_1}\sum_{j=1}^{\beta}\log\left(-\frac{q^j(t,\ u(t))}{M}\right)dt,\quad u(\cdot)\in B_2,$$

取
$$p_i'(u)=\varepsilon_i p'(u),\quad i=0,\ 1,\ \cdots,\ u\in B_2,$$

$\varepsilon_i\to 0$; 可验证得知 $\{p_i'(u)\}$ 是关于集 B_2 的内罚函数列.

对于约束(9)和(11),作泛函

$$p''(u)=\int_{t_0}^{t_1}\sum_{j=1}^{\gamma}(r^j(t,\ x(t,\ u(\cdot))))^2\,dt+\sum_{j=1}^{\alpha}(g^j(x(t_1,\ u(\cdot))))^2,$$

取
$$p_i''(u)=\frac{1}{\varepsilon_i}\,p''(u),\quad i=0,\ 1,\ \cdots,$$

$\varepsilon_i\to 0$, 可验证 $\{p_i''(u)\}$ 是关于集 $B_1\cap B_3$ 的外罚函数列.

同样地,对约束(12)和(13),可以类似于对约束(10)的处理方法而分别地构造出关于集 B_4 和集 B_5 的外罚函数和内罚函数; 若要求同时满足约束(12)和(13),则可直接构造出关于集 $B_4\cap B_5$ 的外罚函数和内罚函数. 例如,外罚函数可取为

$$p_i''(u)=\frac{1}{\varepsilon_i}\int_{t_0}^{t_1}\left\{\sum_{j=1}^{\mu}(\max\{0,\ s^j(t,\ x(t,\ u))\})^2\right.$$
$$\left.+\sum_{k=1}^{\nu}(\max\{0,\ w^k(t,\ x(t,\ u),\ u(t))\})^2\right\}dt,$$

$\varepsilon_i>0$ 且 $\varepsilon_i\to 0$; 内罚函数可取为

$$p_i'(u)=-\varepsilon_i\int_{t_0}^{t_1}\sum_{j=1}^{\mu}\log(-s^j(t,\ x(t,\ u)))$$
$$+\sum_{k=1}^{\nu}\log(-w^k(t,\ x(t,\ u),\ u(t)))\}dt,$$

此处 $\varepsilon_i>0$ 且 $\varepsilon_i\to 0$, 这里仍假设 $s^j(t,\ x)$ 和 $w^k(t,\ x,\ u)$, $t\in\mathbb{R}$, $x\in\mathbb{R}^n$, $u\in\mathbb{R}^m$ 是下方有界的,在上式中之所以省略这一界限,是因为它并不影响求最优解.

上述的罚函数方法是建立在由 $(\mathrm{UPI})_i$ 或 $(\mathrm{UPE})_i$, $i=0,\ 1,\ 2,\ \cdots$, 的解序列 $\{u_{i*}\}$ 收敛于(CP)之解的基础上的. 这事实上是依赖于对无约束最优化问题 $(\mathrm{UPI})_i$ 或 $(\mathrm{UPE})_i$, $i=0,\ 1,\ 2,\ \cdots$,

的解，能用无约束最优化计算方法例如§1、§2的一些方法计算出来. 但各种无约束最优化计算方法多是仅求出梯度的零点，且对每个 $(UPI)_i$ 或 $(UPE)_i$ 都要作无限次迭代才能求出 u_{i*}，因此，在实践上要想在短时间内完成对 (CP) 求解是很困难的. 故人们提出只作有限次迭代以求出 $(UPI)_i$ 或 $(UPE)_i$ 的近似解的算法. 下面就是这种想法的求解最优控制问题(1)～(7)或即最优化问题(CP)的内罚函数型的一种算法. 关于外罚函数型的算法，留给读者讨论.

取

$$J_i(u(\cdot))=J(u(\cdot))+\varepsilon_i p'(u),$$

此处 $J(u)$ 见式(2)，$\varepsilon_i=\varepsilon/\beta^i$，$\varepsilon>0$，$\beta>1$；$\varepsilon_i p'(u)$，$i=0, 1, 2, \cdots$，是关于集 B 的内罚函数列，例如取前面就(9)～(13)的约束所构造的内罚函数，其形式为

$$p'(u)=\int_{t_0}^{t_1}\bar{p}'(t, x(t), u(t))dt,$$

$\bar{p}'(\cdot)$ 与 $f^0(\cdot)$ 有相同的连续性质. 设 $\mathring{J}=J_*$ 及假设 $\mathring{H}$ 成立，由此引理4已指出集 $C_i=\{u|J_i(u)\leqslant J(u'), u\in\mathring{B}\}$ 非空且紧致，为简单起见，认为 $i=0, 1, 2, \cdots$.

算法1 (内罚函数型的 Goldstein 梯度法)

第0步 选取 $\varepsilon>0$，$\alpha\in\left(0, \dfrac{1}{2}\right)$，$\beta>1$，

$$\bar{v}_0(\cdot)\in C_0.$$

第1步 置 $\varepsilon_0=\varepsilon$，置 $j=0$，置 $i=0$.

第2步 取 $u(t)=\bar{u}_j(t)$，$t\in[t_0, t_1]$，解微分方程(1)而得

$$\bar{x}_j(t)=x(t, \bar{u}_j(\cdot)),\quad t\in[t_0, t_1].$$

第3步 在 $u(t)=\bar{u}_j(t)$，$x(t)=\bar{x}_j(t)$，$t\in[t_0, t_1]$ 之下求微分方程组

$$\frac{d}{dt}\psi(t)=-\frac{\partial f(t,\ x(t),\ u(t))}{\partial x}\psi(t)+\frac{\partial f^0(t,\ x(t),\ u(t))}{\partial x}$$
$$+\varepsilon_i\frac{\partial\bar{p}'(t,\ x(t),\ u(t))}{\partial x}$$

的终值 $\psi(t_1)=0$ 的特解 $\bar{\psi}_j(t)$, $t\in[t_0,\ t_1]$.

第 5 步　置

$$h(\bar{u}_j(\cdot),\ \varepsilon_i)=-\text{grad}\quad J_i(\bar{v}_j(\cdot))(t);$$

若 $\|h(\bar{u}_j,\ \varepsilon_i)\|_2\leqslant\varepsilon_i$, 置 $\varepsilon_{i+1}=\varepsilon_i/\beta$, 置 $u_i(t)=\bar{v}_j(t)$, 置 $i=i+1$, 转第 3 步；否则到第 6 步.

第 6 步　计算数 $\bar{\lambda}_j$, 适合

$$-\bar{\lambda}_j(1-\alpha)\|\text{grad}\,J_i(\bar{u}_j)\|_2^2$$
$$\leqslant\theta(\bar{\lambda}_j,\ \bar{u}_j,\ \varepsilon_i)\leqslant-\bar{\lambda}_j\alpha\|\text{grad}\,J_i(\bar{u}_j)\|_2^2$$

此处

$$\theta(\bar{\lambda}_j,\ \bar{v}_j,\ \varepsilon_i)=J_i(\bar{u}_j+\bar{\lambda}_jh(\bar{u}_j,\ \varepsilon_i))-J_i(\bar{u}_j),$$
$$J_i(\bar{v}_j)=J(\bar{u}_j)+\varepsilon_i\bar{p}'(u_j).$$

第 7 步　置 $\bar{u}_{j+1}=\bar{u}_j+\bar{\lambda}_jh(\bar{u}_j,\ \varepsilon_i)$, 置 $j=j+1$, 并转第 2 步.

以上算法第 2 步至第 5 步实际上是寻找 $J_i(u)$ 的极小点，或求 $h(u,\ \varepsilon_i)$ 的零点（可与 § 2 的算法 4 相比较）. 关于由上述算法所构造出的序列 $\{u_i\}$ 的收敛性问题，我们有下面的

引理 7　算法 1 所构造出的无限序列 $\{u_i\}$, 必致使

$$\lim_{i\to+\infty}\|h(u_i,\ \varepsilon_i)\|_2=0. \tag{23}$$

证　易见有关系

$$C_0\subset C_1\subset\cdots\subset Z.$$

由此可见，对 $i=-1,\ 0,\ 1,\ 2,\ \cdots,\ u_{-1}=\bar{u}_0,\ C_{-1}=C_0$ 有

$$D_i(u_{i-1})\triangleq\{u|J(u)+\varepsilon_i\bar{p}'(u)$$
$$\leqslant J(u_{i-1})+\varepsilon_i\bar{p}'(u_{i-1}),\ u\in\mathring{B},\ u_{i-1}\in C_{i-1}\}\subset C_i,$$
$$i=0,\ 1,\ 2,\ \cdots,$$

从而 $D_i(u_{i-1})$ 是紧致集. 这样, $D_i(u_{i-1})$ 相当于§2的定理2中的集 $C(u_0)$; 因此, 由该定理得知, 算法1构造出一个序列 $\{\bar{u}_k\}$, $k=j, j+1, \cdots, \bar{u}_k\in D_i(u_{i-1})$, 这个序列包含一个收敛子序列, 仍记为 $\{\bar{u}_k\}$, 使 $\|h(\bar{u}_k, \varepsilon_i)\|_2\to 0$, 当 $k\to+\infty$. 故存在整数 $N(i)\geqslant 0$, 使 $\|h(\bar{u}_{j+N(i)}, \varepsilon_i)\|\leqslant\varepsilon_i$, 所以, 按算法1第5步知 $u_i=\bar{u}_{j+N(i)}$, 从而

$$\|h(u_i, \varepsilon_i)\|\leqslant\varepsilon_i.$$

因为 $i\to+\infty$ 时 $\varepsilon_i\to 0$, 故得式(23)成立. ∎

这引理说明了, 并不要求 u_i 使 $h(u_i, \varepsilon_i)=0$, 而只要(23)式成立, 从而这样也就能求得当 $i\to+\infty$ 时 $h(u, \varepsilon_i)$ 的零点的近似值了.

习　题

1　对最优控制问题(1)、(2), 试分别求出约束形式为

$$a_j\leqslant u_j(t)\leqslant b_j, \quad j=1, 2, \cdots, m, \quad t\in[t_0, t_1];$$

$$\int_{t_0}^{t_1}M(t, x(t))dt\leqslant M_0; \quad \int_{t_0}^{t_1}N(t, u(t))dt\leqslant N_0$$

时的相应的内罚函数列和外罚函数列, 此处 a_j、b_j 和 M_0、N_0 是常数, $M(t,x)$ 和 $N(t, u)$ 是其变元的连续函数.

2　在 Banach 空间 L^∞ 内考虑约束最优化问题(P): $\inf\{J(u)\,|\,u\in B\}$, $B\triangleq\{u\,|\,u\in L^\infty, r(u)=0, s(u)\leqslant 0\}$, 此处 $J:L^\infty\to\mathbb{R}$, $r:L^\infty\to\mathbb{R}^\gamma$, $s:L^\infty\to\mathbb{R}^\mu$ 均为其变元的连续函数; 令 $\{p_i(u)\}$ 是关于集 $D\triangleq\{u\,|\,u\in L^\infty, r(u)=0\}$ 的外罚函数; 考虑问题(P_i): $\inf\{J_i(u)\,|\,u\in L^\infty, s(u)\leqslant 0\}$, $i=1, 2, \cdots$, 此处 $J_i(u)\triangleq J(u)+p_i(u)$. 求证: 在某些充分条件下, 问题($P_i$)有解且($P_i$), $i=1, 2, \cdots$, 的解序列 $\{u_{i*}\}$ 的任一个聚点均为问题(P)的解.

3　试分别在下面的假设条件下作出相应的关于集 B 的外罚函数型的 Goldstein 梯度算法, 并指出相应于引理7的收敛定理为真:

(1) 设集 $\{u\,|\,J(u)\leqslant a\}$ 对任定的 $a\in\mathbb{R}$ 均为紧致集.

(2) 设本节初所作的假设 H 成立.

第六章 分布参数系统的最优控制

§1 引 言

如果受控系统的动态方程是微分差分方程、积分微分方程或偏微分方程，那末系统的状态空间是无限维的，而不是有限维的．这种状态空间为无限维的系统，称为*分布参数系统*．前面所讨论的用常微分方程组描述的系统，称为*集中参数系统*．

自从集中参数系统的最优控制理论建立后，人们期望把它推广到分布参数系统．

苏联学者布特可夫斯基(А. Г. Бутковский)在讨论炉温控制时，首先把热传导方程的某种最优控制问题转化为庞特里雅金讨论过的问题，开创了分布参数系统最优控制理论的研究．在美国，华人学者王耿介(P. K. C. Wang)，联系航天技术中的控制问题，于1964年系统地讨论了分布参数系统的控制问题．法托里尼(H. O. Fattorini)则讨论了巴拿赫空间中发展方程的时间最优控制问题．在法国，科学院院士利翁斯(J. L. Lions)利用变分不等式，系统地研究了分布参数系统的二次最优控制理论，并于1968年出版了专著．在我国，联系飞行器的镇定问题，关肇直、宋健等研究了分布参数系统的镇定问题；张学铭研究了某些波动方程的最优控制问题．最近，宋健等利用分布参数系统理论研究了中国的人口预测与控制问题．

在分布参数系统最优控制理论的研究中，人们利用凸分析方法，因此假设控制区域 U 是凸闭集．这与集中参数系统的最优控制理论对控制区域 U 没有附加要求有明显的不同．人们自然要

问：在分布参数系统最优控制理论中，能否去掉关于 U 的凸性要求？李训经、姚允龙等讨论过该问题．他们讨论了时间最优控制问题和最大值原理．

本章介绍分布参数系统最优控制理论初步，主要是线性二次最优控制问题和时间最优控制问题．

显然，为研究分布参数系统最优控制问题，泛函分析初步[特别是希尔伯特(D. Hilbert)空间理论]以及偏微分方程初步[特别是拉普拉斯(P. S. Laplace)方程、热传导方程和波动方程]等基础知识是必需的．另外，巴拿赫(S. Banach)空间的凸集，算子半群理论，索伯列夫(С. Л. Соболев)空间的理论等，也是需要的．因此，初学者可以不读这一章．

§2 希尔伯特空间上二次泛函的极值问题

设 H 是一希尔伯特空间，即它是一个线性空间，并且对于 H 中的一对元素 x 和 y，存在实数 $\langle x, y\rangle$ 与之对应，它称为内积，并适合下述条件：

1° $\langle x, x\rangle \geqslant 0$；当且仅当 $x=0$ 时 $\langle x, x\rangle=0$；

2° $\langle x, y\rangle=\langle y, x\rangle$；

3° 对于 $\alpha_1, \alpha_2\in\mathbb{R}$ 和 $x_1, x_2, y\in H$，成立着

$$\langle \alpha_1x_1+\alpha_2x_2, y\rangle=\alpha_1\langle x_1, y\rangle+\alpha_2\langle x_2, y\rangle;$$

4° 在范数 $\|x\|=(\langle x, x\rangle)^{\frac{1}{2}}$ 下，H 是完备空间，即如果

$$\lim_{n,m\to\infty}\|x_n-x_m\|=0,$$

那末存在 $x\in H$，使得

$$\lim_{n\to\infty}\|x_n-x\|=0.$$

设 U 是 H 中的凸闭集，即对于 $x_1, x_2\in U$，$\alpha\in[0, 1]$，成立着

$$\alpha x_1+(1-\alpha)x_2\in U,$$

并且如果 $x_k\in U$, $\lim\limits_{k\to\infty}x_k=x$, 那末 $x\in U$.

设 $\pi: H\times H\to\mathbb{R}$ 具有性质:

i) 对称性: $\pi(x, y)=\pi(y, x)$;

ii) 线性: $\pi(\alpha_1x_1+\alpha_2x_2, y)=\alpha_1\pi(x_1, y)+\alpha_2\pi(x_2, y)$;

iii) 连续性: 存在常数 $M>0$, 使得

$$|\pi(x, y)|\leqslant M\|x\|\cdot\|y\|;$$

iv) 强制正性: 存在常数 $c>0$ 使得

$$\pi(x, x)\geqslant c\|x\|^2.$$

具有上述性质的 π 称为对称双线性泛函, 并且它是连续的和强制正的.

连续线性泛函 $L: H\to\mathbb{R}$ 具有性质:

$$L(\alpha_1x_1+\alpha_2x_2)=\alpha_1L(x_1)+\alpha_2L(x_2);$$

$$|L(x)|\leqslant M_1\|x\|,$$

这里 α_1, $\alpha_2\in\mathbb{R}$, x, x_1, $x_2\in H$, 常数 $M_1>0$.

二次泛函的极值问题 在 U 上讨论

$$J(x)=\pi(x, x)-2L(x)$$

的极小值问题, 即求 $x_0\in U$ 使得

$$J(x_0)\leqslant J(x)$$

对于 $x\in U$ 都成立.

定理 1 设 U 是希尔伯特空间中的凸闭集, π 是 H 上强制正的连续对称双线性泛函, L 是 H 上的连续线性泛函, 那末存在唯一的 $x_0\in U$, 使得

$$J(x_0)\leqslant J(x)$$

对于 $x\in U$ 都成立.

证 因为

$$J(x)\geqslant c\|x\|^2-M_1\|x\|=\|x\|(c\|x\|-M_1),$$

所以

$$\alpha=\inf_{x\in U}J(x)\geqslant\inf_{x\in U}\|x\|(c\|x\|-M_1)>-\infty.$$

设$\{x_k\}$是上述问题的极小化序列，即

$$x_k\in U,\quad \lim_{k\to\infty}J(x_k)=\alpha.$$

根据π的双线性性和L的线性性，成立着平行四边形法则，即有

$$J\left(\frac{x_k+x_l}{2}\right)+\pi\left(\frac{x_k-x_l}{2},\ \frac{x_k-x_l}{2}\right)=\frac{1}{2}\{J(x_k)+J(x_l)\}.$$

所以

$$\frac{c}{4}\|x_k-x_l\|^2\leqslant\frac{1}{2}\{J(x_k)+J(x_l)\}-J\left(\frac{x_k+x_l}{2}\right).$$

因为x_k，$x_l\in U$和U的凸性，知$(x_k+x_l)/2\in U$，所以

$$J\left(\frac{x_k+x_l}{2}\right)\geqslant\alpha.$$

从而

$$0\leqslant\frac{c}{4}\|x_k-x_l\|^2\leqslant\frac{1}{2}\{J(x_k)+J(x_l)\}-\alpha.$$

当k，$l\to\infty$时，上式右端趋于0，所以$\{x_k\}$是柯西序列，根据H的完备性和U的闭性，存在$x_0\in U$，使得

$$\lim_{k\to\infty}x_k=x_0.$$

又因为

$$\begin{aligned}|J(x_k)-J(x_0)|&\leqslant|\pi(x_k,\ x_k)-\pi(x_0,\ x_0)|+2|L(x_k-x_0)|\\&\leqslant|\pi(x_k,\ x_k-x_0)|+|\pi(x_k-x_0,\ x_0)|+2M_1\|x_k-x_0\|\\&\leqslant(M(\|x_k\|+\|x_0\|)+2M_1)\|x_k-x_0\|.\end{aligned}$$

所以

$$J(x_0)=\lim_{k\to\infty}J(x_k)=\alpha.$$

即证明了$x_0\in U$使$J(x_0)=\alpha$的存在性。

再证唯一性. 如果$\bar{x}_0\in U$也满足

$$J(\bar{x}_0)=\alpha,$$

由平行四边形法则，成立着

$$J\left(\frac{x_0+\bar{x}_0}{2}\right)+\pi\left(\frac{x_0-\bar{x}_0}{2}, \frac{x_0-\bar{x}_0}{2}\right)$$

$$=\frac{1}{2}\{J(x_0)+J(\bar{x}_0)\}=\alpha$$

所以 $$\frac{c}{4}\|x_0-\bar{x}_0\|^2\leqslant 0,$$

即 $$\bar{x}_0=x_0.$$ ▌

定理 2 在定理 1 的条件下，$J(x)$ 在凸闭集 U 上的极小点 x_0 由下面的变分不等式表征：对于 $x\in U$ 成立着

$$\pi(x_0, x-x_0)\geqslant L(x-x_0). \tag{1}$$

证 设 $x_0\in U$ 是 $J(\cdot)$ 的极小点，$x\in U$，那末当 $\alpha\in(0, 1)$ 时 $x_0+\alpha(x-x_0)\in U$，所以

$$\begin{aligned}J(x_0)&\leqslant J(x_0+\alpha(x-x_0))\\&=J(x_0)+2\alpha\pi(x_0, x-x_0)+\alpha^2\pi(x-x_0, x-x_0)\\&\quad-2\alpha L(x-x_0).\end{aligned}$$

从而

$$0\leqslant 2\alpha\{\pi(x_0, x-x_0)-L(x-x_0)\}+\alpha^2\pi(x-x_0, x-x_0).$$

两端除以 2α，并令 $\alpha\to 0$，得到(1)。

反之，如果(1)对于 $x\in U$ 成立。因为

$$\pi(x-x_0, x-x_0)\geqslant c\|x-x_0\|^2\geqslant 0,$$

所以

$$\begin{aligned}0&\leqslant 2\{\pi(x_0, x-x_0)-L(x-x_0)\}+\pi(x-x_0, x-x_0)\\&=\pi(x, x)-2L(x)-\pi(x_0, x_0)+2L(x_0).\end{aligned}$$

即 $$J(x_0)\leqslant J(x)$$

对于 $x\in U$ 都成立。所以 $x_0\in U$ 是 $J(\cdot)$ 的极小点。 ▌

不等式(1)称为变分不等式。

根据定理 1 和定理 2 推知，变分不等式(1)在 U 上存在唯一

的解 x_0.

注　如果在 H 中利用双线性泛函 π 来定义内积 $[x, y]=\pi(x, y)$，根据 i)～iv)，$[x, y]$ 满足内积的性质 1°～3°. 若记 $x\in H$ 的范数 $|x|$：

$$|x|^2=[x, x]=\pi(x, x).$$

那末
$$c\|x\|^2\leqslant|x|^2\leqslant M\|x\|^2.$$

所以，在范数 $|\cdot|$ 下，H 是完备的希尔伯特空间.

根据希尔伯特空间上连续线性泛函的黎斯(F. Riesz)表示定理，存在 $\bar{x}\in H$，使得

$$L(x)=[\bar{x}, x].$$

这样

$$\begin{aligned}J(x)&=[x, x]-2[\bar{x}, x]\\&=[x-\bar{x}, x-\bar{x}]-[\bar{x}, \bar{x}]=|x-\bar{x}|^2-|\bar{x}|^2.\end{aligned}$$

所以求 $J(\cdot)$ 在 U 上的极小点 x_0，等价于在 U 上求 x_0，使得它与 $\bar{x}$ 的距离 $|x_0-\bar{x}|$ 为最小值，即

$$|x_0-\bar{x}|\leqslant|x-\bar{x}|$$

对于 $x\in U$ 成立. 因此，可以用第二章 §1 的方法. 事实上，定理 1 和定理 2 的证明方法也和那里类似.

系 1　如果 $U=H$，那末 $J(\cdot)$ 的极小点 x_0 适合欧拉方程

$$\pi(x_0, x)=L(x) \tag{2}$$

对于 $x\in H$ 成立.

证　因为 $U=H$，所以对于 $x\in H$，$\pm x+x_0\in H$，从而

$$\pi(x_0, x)\geqslant L(x),\quad \pi(x_0, -x)\geqslant L(-x).$$

因此，得到(2). ∎

系 2　如果 U 是 H 中的凸锥，那末 $J(\cdot)$ 的极小点 x_0 适合

$$\pi(x_0, x)\geqslant L(x),\quad \forall x\in U, \tag{3}$$

$$\pi(x_0, x_0)=L(x_0). \tag{4}$$

证 因为 U 是凸锥，由 $x_0\in U$ 和 $x\in U$ 推知 $x+x_0\in U$，从而由(1)得到(3).

又因为 $0\in U$，所以由(1)得到

$$\varpi(x_0, -x_0)\geqslant L(-x_0),$$

再由(3)得到

$$\varpi(x_0, x_0)\geqslant L(x_0).$$

从而(4)成立. ∎

例 题

一 索伯列夫空间

设 Ω 是 $\mathbb{R}^n$ 中的有界区域，它的边界 Γ 是 $n-1$ 维光滑流形.

如果 $\varphi:\Omega\to\mathbb{R}$ 是无穷次可微的，它的支集是 $\{x\,|\,\varphi(x)\neq 0\}$ 的闭包. Ω 上的支集含在 Ω 内的无穷次可微函数的全体，记为 $\mathscr{D}(\Omega)$，在通常的加法和数乘的运算下，它是线性空间.

我们称 φ_k 在 $\mathscr{D}(\Omega)$ 中收敛于 φ，如果

1° φ_k 的支集含在与 k 无关的有界闭集 K 中，而 $K\subset\Omega$;

2° φ_k 和它的各阶导数在 Ω 中都一致收敛于 φ 及其各阶导数.

在 $\mathscr{D}(\Omega)$ 上引进上述的极限，相当于引进一种拓扑，使它成为线性拓扑空间.

$\mathscr{D}(\Omega)$ 上的连续线性泛函的全体，称为 $\mathscr{D}(\Omega)$ 的对偶空间，记为 $\mathscr{D}'(\Omega)$.

如果 $f\in L^2(\Omega)$，那末

$$\langle f, \varphi\rangle=\int_\Omega f(x)\varphi(x)dx$$

定义了 $\mathscr{D}(\Omega)$ 上的一个连续线性泛函. 因此，$L^2(\Omega)\subset\mathscr{D}'(\Omega)$.

对于 $f\in\mathscr{D}'(\Omega)$，$\langle f, \cdot\rangle$: $\mathscr{D}(\Omega)\to\mathbb{R}$ 是连续线性泛函. 由于

当 $\varphi\in\mathscr{D}(\Omega)$ 时，$\frac{\partial\varphi}{\partial x_j}\in\mathscr{D}(\Omega)$，从而下式

$$-\left\langle f,\frac{\partial\varphi}{\partial x_j}\right\rangle$$

决定了 $\mathscr{D}(\Omega)$ 上的一个连续线性泛函，记它为

$$\langle f_j,\varphi\rangle;$$

我们称 f_j 是 f 关于 x_j 的广义导数，记为 $\frac{\partial f}{\partial x_j}$. 因此，对于 $f\in\mathscr{D}'(\Omega)$，皆有广义导数.

我们已知 $f\in L^2(\Omega)$，那末 $f\in\mathscr{D}'(\Omega)$，从而 $L^2(\Omega)$ 中的函数皆有广义导数. 我们又知道，存在处处连续处处不可微的函数. 所以，广义导数的存在并不保证导数的存在.

如果对于 $f\in\mathscr{D}'(\Omega)$，存在 $\tilde{f}(\cdot)\in L^2(\Omega)$，使得

$$\langle f,\varphi\rangle=\int_\Omega\tilde{f}(x)\varphi(x)dx$$

对于 $\varphi\in\mathscr{D}(\Omega)$ 皆成立，那末我们就认为 f 是 $\tilde{f}(\cdot)$，从而可以视 $\mathscr{D}'(\Omega)$ 上的元素 f 为 Ω 上的函数. 并且，认为 $f\in L^2(\Omega)$.

如果 $f\in\mathscr{D}'(\Omega)$ 有直到 m 阶平方可积的广义导数，就称 $f\in H^m(\Omega)$，且记 f 的范数 $\|f\|_m$ 为

$$\|f\|_m^2=\int_\Omega\sum_{|\alpha|\le m}|D^\alpha f|^2\,dx,\tag{5}$$

其中

$$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n),\ |\alpha|=\alpha_1+\alpha_2+\cdots+\alpha_n,$$

$$D^\alpha f=\frac{\partial^{|\alpha|}f}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}}.$$

可以证明：在上述范数下 $H^m(\Omega)$ 是完备的，并且范数满足平行四边形法则，从而它是一个希尔伯特空间. 该空间的内积是

$$\langle f,\tilde{f}\rangle_m=\int_\Omega\sum_{|\alpha|\le m}(D^\alpha f)(D^\alpha\tilde{f})dx.$$

空间 $H^m(\Omega)$ 称为索伯列夫空间. 显然，如果 f 在 $\overline{\Omega}$ 上有 m

阶连续偏导数，那末$f\in H^m(\Omega)$. 这时，f在Γ上的值是完全确定的.

对于在$\bar{\Omega}$上有m阶连续偏导数的函数f，可以用等式(5)定义f的范数$\|f\|_m$，可以证明，具有这种性质的函数在范数(5)下完备化得到的空间就是$H^m(\Omega)$.

同样，由于Γ是$n-1$维的光滑流形，并且是有界的，我们称在Γ上无穷次可微的函数全体构成$\mathscr{D}(\Gamma)$，然后利用Γ的测度可以用上述方法定义$H^m(\Gamma)$.

对于$\Omega=\mathbb{R}^n$，我们设$\mathscr{D}(\mathbb{R}^n)$是具有有界支集的无穷次可微函数的全体，它的元素列收敛可以类似于前面来定义，从而可以定义$\mathscr{D}'(\mathbb{R}^n)$，$H^m(\mathbb{R}^n)$.

对于$f\in L^2(\mathbb{R}^n)$，可以定义它的富里埃(J. Fourier)变换

$$\mathscr{F}f=\int_{\mathbb{R}^n}\exp(2\pi i x\cdot\xi)f(x)\,dx,$$

这里$x\cdot\xi=\sum_{j=1}^n x_j\xi_j$. 而且

$$\|\mathscr{F}f\|^2=\|f\|^2=\int_{\mathbb{R}^n}|f(x)|^2dx.$$

显然成立着

$$\begin{aligned}\mathscr{F}\left(\frac{\partial f}{\partial x_j}\right)&=\int_{\mathbb{R}^n}\exp(2\pi i x\cdot\xi)\frac{\partial f(x)}{\partial x_j}dx\\&=-\int_{\mathbb{R}^n}(2\pi i\xi_j)\exp(2\pi i x\cdot\xi)f(x)dx\\&=-2\pi i\xi_j\mathscr{F}f.\end{aligned}$$

因此，$f\in H^m(\mathbb{R}^n)$等价于

$$\xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_n^{\alpha_n}\mathscr{F}f\ (\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n),\ |\alpha|\leqslant m)$$

在$\mathbb{R}^n$上平方可积，从而等价于

$$(1+|\xi|^2)^{\frac{m}{2}}\mathscr{F}f\in L^2(\mathbb{R}^n).$$

今后，如果$(1+|\xi|^2)^{\frac{m}{2}}\mathscr{F}f\in L^2(\mathbb{R}^n)$，就称$f\in H^m(\mathbb{R}^n)$.

对于实数 s，如果 $(1+|\xi|^2)^{\frac{s}{2}}\mathscr{F}f\in L^2(\mathbb{R}^n)$，就称 $f\in H^s(\mathbb{R}^n)$。这就定义了分数次的索伯列夫空间 $H^s(\mathbb{R}^n)$。

用适当的方法也可以定义 $H^s(\Gamma)$。

空间

$$H^{m-\frac{1}{2}}(\Gamma)\times H^{m-\frac{3}{2}}(\Gamma)\times\cdots\times H^{\frac{1}{2}}(\Gamma)$$

称为 $H^m(\Omega)$ 的迹空间。

可以证明：存在线性有界算子 γ 映 $H^m(\Omega)$ 到它的整个迹空间，使得当 $f\in C^m(\overline{\Omega})$（$\overline{\Omega}$ 上的 m 阶连续可微函数的全体）时

$$\gamma f=\left(f\Big|_\Gamma,\ \frac{\partial f}{\partial n}\Big|_\Gamma,\ \cdots,\ \frac{\partial^{m-1}f}{\partial n^{m-1}}\Big|_\Gamma\right),$$

这里 n 是 Γ 的外法线方向。所以，当 $f\in H^m(\Omega)$ 时，我们称 γf 是 f 的边界值。

空间

$$H_0^m(\Omega)=\{f\,|\,\gamma f=0,\ f\in H^m(\Omega)\}$$

是 $H^m(\Omega)$ 的子空间。

可以证明：$\mathscr{D}(\Omega)$ 在范数(5)下完备化得到 $H_0^m(\Omega)$。

显然成立着

$$\mathscr{D}(\Omega)\subset H_0^m(\Omega)\subset H^m(\Omega)\subset L^2(\Omega)\subset\mathscr{D}'(\Omega).$$

如果把 $L^2(\Omega)$ 的对偶空间与 $L^2(\Omega)$ 等同，那末 $H_0^m(\Omega)$ 的对偶空间称为 $H^{-m}(\Omega)$，即

$$(H_0^m(\Omega))'=H^{-m}(\Omega).$$

对于 $H^s(\Gamma)$，成立着

$$(H^s(\Gamma))'=H^{-s}(\Gamma).$$

这些结论的详细讨论，请查阅有关索伯列夫空间的专著，这里省略了。

二　狄里克莱问题

设 $a_0(\cdot)$、$a_{jk}(\cdot)\in C^1(\overline{\Omega})$，满足条件

$$a_0(x) \geqslant \alpha > 0, \quad x \in \Omega,$$

$$\sum_{j,k=1}^{n} a_{jk}(x)\xi_j\xi_k \geqslant \alpha(\xi_1^2+\xi_2^2+\cdots+\xi_n^2),\ x\in\Omega, \quad \xi\in\mathbb{R}^n. \tag{6}$$

狄里克莱(Dirichlet)问题是：寻求 $y(\cdot)$ 满足

$$\sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\right)+a_0(x)y(x)=f(x), \quad x\in\Omega, \tag{7}$$

$$y(x)=0, \quad x\in\Gamma, \tag{8}$$

这里 $f(\cdot)\in L^2(\Omega)$.

首先，如果 $y(\cdot)$ 是狄里克莱问题的解，用 $\tilde{y}(\cdot)\in\mathscr{D}(\Omega)$ 乘 (7) 的两端，并运用格林(Green)公式，得到

$$\int_\Omega\left\{\sum_{j,k=1}^{n} a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\,\frac{\partial \tilde{y}(x)}{\partial x_j}+a_0(x)y(x)\tilde{y}(x)\right\}dx$$

$$=\int_\Omega f(x)\tilde{y}(x)dx. \tag{9}$$

上式对于 $\tilde{y}(\cdot)\in H_0^1(\Omega)$ 也是成立的.

今后，如果 $y(\cdot)\in H_0^1(\Omega)$ 适合(9)，我们称它为狄里克莱问题(7)、(8)的广义解或弱解.

如果对于 $y(\cdot)$, $\tilde{y}(\cdot)\in H_0^1(\Omega)$，定义

$$\pi(y,\ \tilde{y})=\int_\Omega\left\{\sum_{j,k=1}^{n} a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\cdot\frac{\partial \tilde{y}(x)}{\partial x_j}+a_0(x)y(x)\tilde{y}(x)\right\}dx,$$

它是 $H_0^1(\Omega)$ 上的连续双线性泛函，并且由(6)成立着

$$\pi(y,\ y)\geqslant\alpha\int_\Omega\left\{\sum_{j=1}^{n}\left(\frac{\partial y(x)}{\partial x_j}\right)^2+y^2(x)\right\}dx$$

$$=\alpha\|y\|_1^2.$$

所以 π 是 $H_0^1(\Omega)$ 上的强制正的连续双线性泛函.

又设 $L(y)$ 定义如下

$$L(y)=\int_\Omega f(x)y(x)dx,$$

它是 $H_0^1(\Omega)$ 上的连续线性泛函.

因此，在 $H_0^1(\Omega)$ 上讨论二次泛函

$$\int_\Omega \left\{ \sum_{j,k=1}^n a_{jk}(x) \frac{\partial y(x)}{\partial x_k} \frac{\partial y(x)}{\partial x_j} \right.$$
$$\left. + a_0(x) y^2(x) - 2f(x) y(x) \right\} dx \tag{10}$$

的极小值时，它的极值元素 $y(\cdot)\in H_0^1(\Omega)$ 存在唯一，并且欧拉方程

$$\int_\Omega \left\{ \sum_{j,k=1}^n a_{jk}(x) \frac{\partial y(x)}{\partial x_k} \frac{\partial \tilde{y}(x)}{\partial x_j} \right.$$
$$\left. + a_0(x) y(x)\tilde{y}(x) - f(x)\tilde{y}(x) \right\} dx = 0,$$

对于 $\tilde{y}(\cdot)\in H_0^1(\Omega)$ 皆成立．它正是(9)．因此，讨论狄里克莱问题(7)、(8)的广义解问题等价于在 $H_0^1(\Omega)$ 上讨论二次泛函(10)的极小值问题．

如果我们研究狄里克莱问题解 y 对 f 的依赖关系，可记它为 $y(f)$．

根据欧拉方程有（当 $\tilde{y}=y$ 时）

$$\int_\Omega \left\{ \sum_{j,k=1}^n a_{jk}(x) \frac{\partial y}{\partial x_k} \frac{\partial y}{\partial x_j} + a_0(x) y^2(x) \right\} dx$$
$$= \int_\Omega f(x) y(x) dx,$$

上式左端大于等于 $\alpha\|y\|_1^2$．因此

$$\alpha\|y\|_1^2 \leqslant \int_\Omega f(x) y(x) dx \leqslant \|f\|\cdot\|y\| \leqslant \|f\|\|y\|_1,$$

所以

$$\|y\|_1 \leqslant \|f\|/\alpha, \tag{11}$$

即 y 对 f 的依赖关系是线性的，并且是线性有界的．

三　诺依曼问题

求方程(7)满足边值条件

$$\left.\frac{\partial y}{\partial \nu_A}\right|_\Gamma = g(x) \tag{12}$$

的问题，称为诺依曼问题．这里 $g\in H^{-\frac{1}{2}}(\Gamma)$ 是给定的，$\frac{\partial y}{\partial \nu_A}$ 表示

$$\frac{\partial y}{\partial \nu_A}\Big|_\Gamma=\sum_{j,k=1}^{n} a_{jk}(x)\ \frac{\partial y}{\partial x_k}\cos(n,\ x_j)\Big|_\Gamma,$$

其中 $\cos(n,\ x_j)$ 是 Γ 的外法向量与 x_j 轴的方向余弦．

为讨论诺依曼问题，在 $H^1(\Omega)$ 上讨论双线性泛函

$$\pi(\varphi,\ \psi)=\int_\Omega\Big\{\sum_{j,k=1}^{n} a_{jk}(x)\ \frac{\partial \varphi}{\partial x_k}\ \frac{\partial \psi}{\partial x_j}+a_0(x)\varphi(x)\psi(x)\ \Big\}dx,$$

它是连续的和强制正的；以及线性泛函

$$L(\varphi)=\int_\Omega f(x)\varphi(x)dx+\int_\Gamma g(x)\varphi(x)dx,$$

由于 $g\in H^{-\frac{1}{2}}(\Gamma)$ 和 $\varphi\in H^1(\Omega)$ 推出的 $\varphi|_\Gamma\in H^{\frac{1}{2}}(\Gamma)$，得知上式第二项也是连续的，所以 L 是连续线性泛函．

显然，在 $H^1(\Omega)$ 上讨论二次泛函

$$\pi(\varphi,\ \varphi)-2L(\varphi)$$

的极小值问题，它的欧拉方程是

$$\int_\Omega\Big\{\sum_{j,k=1}^{n} a_{jk}(x)\ \frac{\partial y(x)}{\partial x_k}\ \frac{\partial \tilde{y}(x)}{\partial x_j}+a_0(x)y(x)\tilde{y}(x)-f(x)\tilde{y}(x)\Big\}dx-\int_\Gamma g(x)\tilde{y}(x)dx=0$$

对于 $\tilde{y}(\cdot)\in H^1(\Omega)$ 皆成立．

特别，当 $\tilde{y}(\cdot)\in H_0^1(\Omega)$ 时，上式成为(9)．所以可以认为 $y(\cdot)\in H^1(\Omega)$ 满足椭圆型方程(7)．

对上式应用格林公式，并且注意到(7)，就有

$$\int_\Gamma\Big\{\sum_{j,k=1}^{n} a_{jk}(x)\ \frac{\partial y}{\partial x_k}\cos(n,\ x_j)-g(x)\Big\}\tilde{y}(x)dx=0 \qquad (13)$$

对于 $\tilde{y}|_\Gamma\in H^{\frac{1}{2}}(\Gamma)$ 成立．所以可以认为 y 在 Γ 上满足边值条件(12)．

四 单侧边值问题

我们已经看到，在不同的空间中讨论二次泛函的极值，可以得到狄里克莱问题和诺依曼问题的解.

现在 $H^1(\Omega)$ 中的凸闭锥

$$\mathscr{U}=\{y(\cdot)\in H^1(\Omega)\,|\,y|_\Gamma\geqslant 0\}$$

上讨论二次泛函(10)的极小值问题. 我们知道，它的解 $y(\cdot)\in\mathscr{U}$ 存在且唯一，它适合变分不等式：

$$\int_\Omega\left\{\sum_{j,k=1}^n a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\frac{\partial\tilde{y}(x)}{\partial x_j}\right.$$
$$\left.+a_0(x)y(x)\tilde{y}(x)-f(x)\tilde{y}(x)\right\}dx\geqslant 0 \tag{14}$$

对于 $\tilde{y}(\cdot)\in\mathscr{U}$ 皆成立. 并且

$$\int_\Omega\left\{\sum_{j,k=1}^n a_{jk}(x)\frac{\partial y}{\partial x_k}\frac{\partial y}{\partial x_j}\right.$$
$$\left.+a_0(x)y^2(x)-f(x)y(x)\right\}dx=0 \tag{15}$$

因为线性空间 $\mathscr{D}(\Omega)\subset\mathscr{U}$，所以从变分不等式(14)得到

$$\int_\Omega\left\{\sum_{j,k=1}^n a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\cdot\frac{\partial\tilde{y}(x)}{\partial x_j}\right.$$
$$\left.+a_0(x)y(x)\tilde{y}(x)-f(x)\tilde{y}(x)\right\}dx=0$$

对于 $\tilde{y}(\cdot)\in\mathscr{D}(\Omega)$ 成立. 所以，$y(\cdot)$是偏微分方程(7)的广义解.

对于(14)应用格林公式，并注意到(7)，得到

$$\int_\Gamma\sum_{j,k=1}^n a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\cos(n,\ x_j)\tilde{y}(x)dx\geqslant 0$$

对于 $\tilde{y}(\cdot)\in\mathscr{U}$成立. 从而不等式

$$\sum_{j,k=1}^n a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\cos(n,\ x_j)\geqslant 0 \tag{16}$$

对于 $x\in\Gamma$ 成立.

同时，由(15)运用格林公式得到

$$\int_{\Gamma}\sum_{j,k=1}^{n} a_{jk}(x)\frac{\partial y(x)}{\partial x_k}\cos(n, x_j)y(x)dx=0.$$

再根据(16)和 $y(x)|_{\Gamma}\geqslant 0$，得到

$$y(x)\Big|_{\Gamma}\cdot\frac{\partial y}{\partial \nu_A}\Big|_{\Gamma}=0.$$

求方程(7)适合边值条件

$$y(x)\Big|_{\Gamma}\geqslant 0,\ \frac{\partial y}{\partial \nu_A}\Big|_{\Gamma}\geqslant 0,\quad y(x)\cdot\frac{\partial y}{\partial \nu_A}\Big|_{\Gamma}=0$$

的解的问题称为单侧边值问题。

根据上述讨论知道，单侧边值问题在 $H^1(\Omega)$ 中的解等价于在 $\mathscr{U}$ 上讨论二次泛函(10)的极值元素。因此，单侧边值问题的解是存在唯一的。

习　题

1　试讨论狄里克莱边值问题

$$\sum_{j=1}^{n}\frac{\partial^2 y(x)}{\partial x_j^2}=f(x),\ x\in\Omega,$$

$$y(x)|_{\Gamma}=0.$$

2　试讨论诺依曼边值问题

$$\sum_{j=1}^{n}\frac{\partial^2 y(x)}{\partial x_j^2}=f(x),\ x\in\Omega,$$

$$\frac{\partial y}{\partial n}\Big|_{\Gamma}=0,$$

这里 n 是 Γ 的外法线方向。

3　试讨论诺依曼问题的解对 f 的依赖关系。

§3　椭圆型系统的最优控制问题

设 V 和 H 是两个希尔伯特空间。V 的内积和范数用记号 $\langle\ ,\ \rangle$ 和 $\|\cdot\|$ 表示；H 的内积和范数用记号 $(\ ,\ )$ 和 $|\cdot|$ 来表示。

假设线性空间 $V\subset H$，并且当 $v\in V$ 时

$$|v|\leqslant c\|v\|, \tag{1}$$

其中 $c>0$ 是常数，即 V 嵌入 H 是连续的．同时，V 在空间 H 中是稠密的．

例如 § 2 中所讨论的空间 $H^1(\Omega)$ 和 $L^2(\Omega)$．显然，成立着

$$H^1(\Omega)\subset L^2(\Omega),$$

并且当 $v\in H^1(\Omega)$ 时

$$|v|^2=\int_\Omega v^2(x)dx\leqslant\int_\Omega\left\{v^2(x)+\sum_{i=1}^n\left(\frac{\partial v}{\partial x_i}\right)^2\right\}dx=\|v\|^2.$$

所以，我们可以取

$$H=L^2(\Omega),\quad V=H^1(\Omega).$$

同时，可以证明 $H^1(\Omega)$ 在 $L^2(\Omega)$ 中是稠密的．

如果把 H 的对偶空间 H' 与 H 等同，即

$$H'=H,$$

那 V 的对偶空间 V' 与 H 成立着关系

$$V\subset H=H'\subset V'. \tag{2}$$

例如，视 $L^2(\Omega)$ 的对偶空间为 $L^2(\Omega)$，那末 $H_0^1(\Omega)$ 的对偶空间为 $H^{-1}(\Omega)$，并且成立着

$$H_0^1(\Omega)\subset L^2(\Omega)\subset H^{-1}(\Omega).$$

在嵌入关系式(2)中，相应的嵌入都是连续的，并且 V 在 H 中稠密，H 在 V' 中稠密．

对于 V 上的连续双线性泛函 $a(\cdot,\ \cdot):V\times V\to\mathbb{R}$，存在 $Au\in V'$，使得

$$a(u,\ v)=\langle Au,\ v\rangle,$$

这里 $\langle,\rangle$ 表示 V' 和 V 的元素的对偶积．

因为对于 $v\in V$，$a(\cdot,\ v):V\to\mathbb{R}$ 是线性的，所以 A 是 V 到 V' 的线性算子．同时，根据 $a(\cdot,\ \cdot)$ 的连续性，得到

$$|\langle Au,\ v\rangle|=|a(u,\ v)|\leqslant M\|u\|\|v\|,$$

从而

$$\|Au\|_{V'}\leqslant M\|u\|,\tag{3}$$

$\|\ \|_{V'}$ 表示 V' 中的范数. 因此, $A\in\mathscr{L}(V, V')$.

再设 $a(\cdot, \cdot)$ 是强制正的, 即存在 $\alpha>0$ 使得

$$a(v, v)\geqslant\alpha\|v\|^2, \quad \forall v\in V.\tag{4}$$

根据定理 2-2, 有

定理 1 对于 $f\in V'$, 存在唯一的 $y\in V$, 使得

$$a(y, \psi)=\langle f, \psi\rangle, \quad \forall\psi\in V.\tag{5}$$

(5) 可写为

$$\langle Ay, \psi\rangle=\langle f, \psi\rangle, \quad \forall\psi\in V,$$

它等价于

$$Ay=f.\tag{6}$$

假设 $\mathscr{U}$ 和 $\mathscr{H}$ 是希尔伯特空间, $B\in\mathscr{L}(\mathscr{U}, V')$, $C\in\mathscr{L}(V, \mathscr{H})$. 又设 $z_d\in\mathscr{H}$ 给定, $N\in\mathscr{L}(\mathscr{U}, \mathscr{U})$ 是自共轭的, 并且存在 $\gamma>0$ 使得

$$\langle Nv, v\rangle_{\mathscr{U}}\geqslant\gamma\|v\|^2_{\mathscr{U}},\tag{7}$$

这里 $\langle, \rangle_{\mathscr{U}}$ 和 $\|\ \|_{\mathscr{U}}$ 表示空间 $\mathscr{U}$ 的内积和范数.

最优控制问题 设 $\mathscr{U}_{ad}$ 是 $\mathscr{U}$ 中的凸闭集, 对于 $v\in\mathscr{U}_{ad}$, 存在唯一的 $y(v)\in V$, 满足

$$Ay(v)=f+Bv,\tag{8}$$

考虑性能指标

$$J(v)=\|Cy(v)-z_d\|^2_{\mathscr{H}}+\langle Nv, v\rangle_{\mathscr{U}}.\tag{9}$$

最优控制问题是选取 $u\in\mathscr{U}_{ad}$, 满足

$$J(u)=\inf_{v\in\mathscr{U}_{ad}} J(v).\tag{10}$$

根据 (9), 有

$$J(v)=\langle C(y(v)-y(0)), C(y(v)-y(0))\rangle_{\mathscr{H}}+\langle Nv, v\rangle_{\mathscr{U}}$$
$$-2\langle z_d-Cy(0), C(y(v)-y(0))\rangle_{\mathscr{H}}+\|z_d-Cy(0)\|^2_{\mathscr{H}}.$$

由于

$$A(y(v)-y(0))=Bv, \tag{11}$$

所以

$$\pi(u,\ v)=\langle C(y(u)-y(0)),\ C(y(v)-y(0))\rangle_{\mathscr{H}}$$
$$+\langle Nu,\ v\rangle_{\mathscr{U}}$$

是 $\mathscr{U}$ 上的连续双线性泛函，并且是强制正的：

$$\pi(v,\ v)\geqslant\gamma\|v\|^2_{\mathscr{U}};$$

而
$$L(v)=\langle z_d-Cy(0),\ C(y(v)-y(0))\rangle_{\mathscr{H}}$$

是 $\mathscr{U}$ 上的连续线性泛函.

定理 2　存在唯一的 $u\in\mathscr{U}_{ad}$ 满足关系(10)，并且它由变分不等式

$$\langle Cy(u)-z_d,\ C(y(v)-y(u))\rangle_{\mathscr{H}}+\langle Nu,\ v-u\rangle_{\mathscr{U}}\geqslant 0 \tag{12}$$

对于 $v\in\mathscr{U}_{ad}$ 皆成立来表征.

证　根据定理 2–1 得到极值元素 $u\in\mathscr{U}_{ad}$ 的存在唯一性. 根据定理 2–2，极值元素 $u\in\mathscr{U}_{ad}$ 由变分不等式

$$\langle C(y(u)-y(0)),\ Cy(v)-Cy(u)\rangle_{\mathscr{H}}+\langle Nu,\ v-u\rangle_{\mathscr{U}}$$
$$\geqslant\langle z_d-Cy(0),\ Cy(v)-Cy(u)\rangle_{\mathscr{H}},\quad \forall v\in\mathscr{U}_{ad}$$

来表征. 上式就是变分不等式(12). ∎

为把变分不等式(12)和状态方程(8)建立明显的联系，我们再来详细讨论(12).

设 $\mathscr{H}'$ 是 $\mathscr{H}$ 的对偶空间，对于 $x\in\mathscr{H}$，$\langle x,\ \cdot\rangle_{\mathscr{H}}$ 决定了 $\mathscr{H}$ 上的线性泛函，所以存在 $\Lambda x\in\mathscr{H}'$，使得 $\langle\Lambda x,\ \cdot\rangle=\langle x,\ \cdot\rangle_{\mathscr{H}}$. 这里 $\langle,\rangle$ 是 $\mathscr{H}'$ 和 $\mathscr{H}$ 中元素的对偶积. $\Lambda:\mathscr{H}\to\mathscr{H}'$ 是线性有界的，称为 $\mathscr{H}$ 到 $\mathscr{H}'$ 的典则同构对应.

若记 $C^*\in\mathscr{L}(\mathscr{H}',\ V')$ 为 $C\in\mathscr{L}(V,\ \mathscr{H})$ 的共轭算子，那末变分不等式(12)可改写为

$$\langle C^*\Lambda(Cy(u)-z_d),\ y(v)-y(u)\rangle+\langle Nu,\ v-u\rangle_{\mathscr{U}}\geqslant 0. \tag{13}$$

设 $p(u)\in V$ 适合

$$A^*p(u)=C^*\Lambda(Cy(u)-z_d),\tag{14}$$

这里 $A^*\in\mathscr{L}(V,V')$ 是 $A\in\mathscr{L}(V,V')$ 的共轭算子. 方程 (14) 的解的存在唯一性,是由 $\langle A^*\varphi,\psi\rangle=\langle\varphi,A\psi\rangle=a(\psi,\varphi)$ 为 V 上的强制正的连续双线性泛函得到的.

由(13)和(14)得到

$$\langle A^*p(u),\ y(v)-y(u)\rangle+\langle Nu,\ v-u\rangle_{\mathscr{U}}\geqslant 0,$$

但

$$\langle A^*p(u),\ y(v)-y(u)\rangle=\langle p(u),\ Ay(v)-Ay(u)\rangle$$
$$=\langle p(u),\ B(v-u)\rangle=\langle B^*p(u),\ v-u\rangle,$$

这里 $B^*\in\mathscr{L}(V,\mathscr{U}')$ 是 $B\in\mathscr{L}(\mathscr{U},V')$ 的共轭算子, $\mathscr{U}'$ 是 $\mathscr{U}$ 的对偶空间. 所以

$$\langle B^*p(u),\ v-u\rangle+\langle Nu,\ v-u\rangle_{\mathscr{U}}\geqslant 0.$$

如果设 $\Lambda_{\mathscr{U}}$ 是 $\mathscr{U}$ 到 $\mathscr{U}'$ 的典则同构对应,那末由上式得到

$$\langle B^*p(u)+\Lambda_{\mathscr{U}}Nu,\ v-u\rangle\geqslant 0,\quad \forall v\in\mathscr{U}_{ad},\quad u\in\mathscr{U}_{ad}.$$

因此证得

定理 3 最优控制问题(10)的解 $u\in\mathscr{U}_{ad}$ 是下述方程和不等式的唯一解:

$$Ay(u)=f+Bu,\tag{8}$$

$$A^*p(u)=C^*\Lambda(Cy(u)-z_d),\tag{14}$$

$$\langle B^*p(u)+\Lambda_{\mathscr{U}}Nu,\ v-u\rangle\geqslant 0,\quad \forall v\in\mathscr{U}_{ad},\tag{15}$$

$$u\in\mathscr{U}_{ad}.$$ ▌

系 1 当 $\mathscr{U}_{ad}=\mathscr{U}$ 时, 最优控制问题 (10) 的唯一解 u 是下述方程的唯一解

$$\begin{aligned}&Ay=f+Bu,\\&A^*p=C^*\Lambda(Cy-z_d),\\&B^*p+\Lambda_{\mathscr{U}}Nu=0.\end{aligned}\tag{16}$$

证 对于 $\mathscr{U}_{ad}=\mathscr{U}$，变分不等式成为

$$\langle B^*p(u)+\Lambda_{\mathscr{U}}Nu,\ v\rangle\geqslant 0,\quad \forall v\in\mathscr{U},$$

从而得到(16)。 ∎

这时，最优的状态 y 由下列方程组决定

$$Ay+BN^{-1}\Lambda_{\mathscr{U}}^{-1}B^*p=f,$$

$$C^*\Lambda Cy-A^*p=C^*\Lambda z_d,$$

而最优控制适合

$$u=-N^{-1}\Lambda_{\mathscr{U}}^{-1}B^*p.$$

例　　题

设 $\mathscr{U}_{ad}$ 是 $L^2(\Omega)$ 中的凸闭集。

$$z_d(\cdot)\in H=L^2(\Omega),\quad f\in L^2(\Omega).$$

对于 $v(\cdot)\in\mathscr{U}_{ad}$，考虑狄里克莱问题

$$\begin{aligned}Ay=-\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y}{\partial x_k}\right)+a_0(x)y=f+v,\quad x\in\Omega,\\ y(x)=0,\quad x\in\Gamma\end{aligned}\tag{17}$$

在 $H_0^1(\Omega)$ 中的唯一解 $y(\cdot,v)$，这里 a_0、a_{jk} 适合 §2 所述的条件。相应于状态 $y(\cdot,v)$ 有泛函

$$J(v)=\int_\Omega(y(x,v)-z_d(x))^2dx+\int_\Omega v^2(x)dx.\tag{18}$$

根据定理 3，如果取

$$V=H_0^1(\Omega),\quad H=L^2(\Omega),$$

$$\mathscr{U}=L^2(\Omega),\quad \mathscr{H}=L^2(\Omega),$$

那末，椭圆型系统的狄里克莱问题(17)关于性能指标(18)的最优控制 $u(\cdot)$，是下述方程和不等式的唯一解：

$$Ay=f+u,\quad y|_\Gamma=0,\tag{19}$$

$$A^*p=y-z_d,\quad p|_\Gamma=0,\tag{20}$$

$$\int_\Omega(p+u)(v-u)dx\geqslant 0,\quad \forall v\in\mathscr{U}_{ad},\tag{21}$$

$$u\in\mathscr{U}_{ad}.$$

当 $\mathscr{U}_{ad}=L^2(\Omega)$ 时，最优状态 y 适合

$$Ay+p=f,\quad x\in\Omega,$$
$$y-A^*p=z_d,\quad x\in\Omega,$$
$$y|_\Gamma=0,\quad p|_\Gamma=0,$$

而最优控制为

$$u=-p,\quad x\in\Omega.$$

* * * * *

对于 $v(\cdot)\in\mathscr{U}_{ad}$，考虑诺依曼问题

$$Ay=-\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y}{\partial x_k}\right)+a_0(x)y$$
$$=f+v,\ x\in\Omega,\tag{22}$$
$$\left.\frac{\partial y}{\partial\nu_A}\right|_\Gamma=g\tag{23}$$

在 $H^1(\Omega)$ 中的解，这里 $g\in H^{-\frac{1}{2}}(\Gamma)$ 给定，$f\in L^2(\Omega)$ 给定. 相应于诺依曼问题的唯一解 $y(\cdot,v)$，称为系统的状态. 性能指标是

$$J(v)=\int_\Omega(y(x,v)-z_d(x))^2dx+\int_\Omega v^2(x)dx.\tag{24}$$

椭圆型系统(22)的诺依曼问题的解关于性能指标(24)的最优控制 $u(\cdot)$，是下述方程和不等式在 $\mathscr{U}_{ad}$ 中的唯一解

$$Ay=f+u,\quad x\in\Omega,$$
$$\left.\frac{\partial y}{\partial\nu_A}\right|_\Gamma=g,$$
$$A^*p=y-z_d,\quad x\in\Omega,$$
$$\left.\frac{\partial p}{\partial\nu_{A^*}}\right|_\Gamma=0,$$
$$\int_\Omega(p(x)+u(x))(v(x)-u(x))dx\geqslant0,\quad\forall v(\cdot)\in\mathscr{U}_{ad},$$
$$u(\cdot)\in\mathscr{U}_{ad}.$$

当 $\mathscr{U}_{ad}=L^2(\Omega)$ 时，最优状态 y 适合

$$Ay+p=f, \quad x\in\Omega,$$
$$y-A^*p=z_d, \quad x\in\Omega,$$
$$\left.\frac{\partial y}{\partial\nu_A}\right|_\Gamma=g, \quad \left.\frac{\partial p}{\partial\nu_{A^*}}\right|_\Gamma=0,$$

最优控制是

$$u=-p.$$

习　题

1 设 $\mathscr{V}=L^2(\Gamma)$，$\mathscr{V}_{ad}$ 是 $\mathscr{V}$ 中的凸闭集．$f\in L^2(\Omega)$，$g\in L^2(\Gamma)$．试讨论下述边界控制问题：系统的状态方程是椭圆型方程

$$-\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y}{\partial x_k}\right)+a_0(x)y=f, \quad x\in\Omega,$$

边值条件为

$$\left.\frac{\partial y}{\partial\nu_A}\right|_\Gamma=g+v, \quad v\in\mathscr{V}_{ad},$$

性能指标为

$$J(v)=\int_\Omega(y(x, v)-z_d(x))^2dx+\int_\Gamma v^2(x)dx,$$

这里 a_0、a_{jk} 满足 §2 的条件，$z_d(\cdot)\in L^2(\Omega)$．

§4　抛物型系统的最优控制问题

设 Ω 是 $\mathbb{R}^n$ 中的有界区域，它的边界 Γ 是 $n-1$ 维光滑流形．$T>0$ 给定，$H=L^2(\Omega)$，$V=H_0^1(\Omega)$．

$\mathscr{U}_{ad}$ 是 $L^2([0, T]; H)$ 中的凸闭集．

设 $f\in L^2([0, T]; H)$，$a_0(\cdot)$、$a_{jk}(\cdot)\in C^1(\bar\Omega)$，满足

$$a_0(x)\geqslant\alpha>0, \quad \forall x\in\Omega,$$
$$\sum_{j,k=1}^{n}a_{jk}(x)\xi_j\xi_k\geqslant\alpha\sum_{j=1}^{n}\xi_j^2, \quad \forall x\in\Omega, \quad \xi\in\mathbb{R}^n,$$

这里 $\alpha>0$ 是常数．

对于 $v\in\mathscr{U}_{ad}$，考虑状态方程为抛物型偏微分方程的系统

$$\frac{\partial y}{\partial t}-\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial y}{\partial x_k}\right)+a_0(x)y=f+v,$$
$$\forall(t,\ x)\in(0,\ T)\times\Omega, \tag{1}$$

它的边值条件为

$$y=0,\quad \forall(t,\ x)\in\Sigma\triangleq(0,\ T)\times\Gamma, \tag{2}$$

初值条件为

$$y(0,\ x)=y_0(x),\quad \forall x\in\Omega, \tag{3}$$

这里 $y_0(\cdot)\in L^2(\Omega)$.

设

$$a(\varphi,\ \psi)=\sum_{j,k=1}^{n}\int_\Omega a_{jk}(x)\frac{\partial\varphi}{\partial x_k}\ \frac{\partial\psi}{\partial x_j}dx$$
$$+\int_\Omega a_0(x)\varphi(x)\psi(x)dx,$$

它是 V 上的强制正的连续双线性泛函. 所以存在 $\mathscr{A}\in\mathscr{L}(V,\ V')$ 使得

$$a(\varphi,\ \psi)=\langle\mathscr{A}\varphi,\ \psi\rangle,\quad \forall\varphi,\ \psi\in V.$$

设
$$D(A)=\{\varphi\,|\,\varphi\in V,\quad \mathscr{A}\varphi\in H\};$$
$$A\varphi=\mathscr{A}\varphi,\qquad \forall\varphi\in D(A).$$

显然, $A:D(A)\to H$ 是线性的, 但是无界算子. 由于 $\mathscr{D}(\Omega)\subset D(A)\subset H$, 所以 $\overline{D(A)}=H$, 即 $D(A)$在 H 中是稠密的. 还可以证明: 如果 $\varphi_k\in D(A)$, $\varphi_k\to\varphi$, $A\varphi_k\to\tilde{\varphi}$ 成立, 那末 $\varphi\in D(A)$, 并且 $\tilde{\varphi}=A\varphi$, 即 A 是闭算子.

可以证明: 存在 H 上的强连续算子半群 $G(t)$, 它以 $-A$ 为母元, 即成立着:

$$G(t)\in\mathscr{L}(H),\qquad \forall t\geqslant0;$$
$$G(0)=I\ (\text{恒等算子});$$
$$G(t+s)=G(t)G(s),\quad \forall t,\ s\geqslant0;$$
$$\lim_{t\to0^+}G(t)\varphi=\varphi,\qquad \forall\varphi\in H;$$

$$\lim_{t\to 0^+}\frac{G(t)\varphi-\varphi}{t}=-A\varphi,\quad \forall\varphi\in D(A).$$

希尔(E. Hille)和吉田耕作讨论了巴拿赫空间上算子半群的理论. 这里是利用该理论讨论方程(1).

记 $y(t)=y(t,\cdot)$, $f(t)=f(t,\cdot)\in H$, $v(t)=v(t,\cdot)\in\mathscr{U}_{ad}$, 那末方程(1)和边值条件(2)可写为希尔伯特空间 H 上的发展方程

$$\frac{dy(t)}{dt}=-Ay(t)+f(t)+v(t),\quad y(t)\in V,\tag{4}$$

而初值条件(3)可写为

$$y(0)=y_0=y_0(\cdot).\tag{5}$$

当 $f=0$, $v=0$ 时, (4)、(5)成为

$$\frac{dy(t)}{dt}=-Ay(t),\quad y(0)=y_0.\tag{6}$$

如果 $y_0\in D(A)$, 那末由

$$\frac{G(t+s)y_0-G(t)y_0}{s}=\frac{G(t)\{G(s)y_0-y_0\}}{s}$$

得知: 当 $s\to 0^+$ 时, 它的极限是存在的; 又由

$$\frac{G(t)y_0-G(t-s)y_0}{s}=G(t-s)\frac{G(s)y_0-y_0}{s}$$

得知, 它当 $s\to 0^+$ 时有极限 $-G(t)Ay_0$. 因此

$$\frac{dG(t)y_0}{d}=-G(t)Ay_0=-AG(t)y_0.$$

所以, $G(t)y_0$是(6)的解, 且以$(0, y_0)$为初值.

对于方程(4)和(5), 我们考虑常数变易公式

$$y(t)=G(t)y_0+\int_0^t G(t-s)\{f(s)+v(s)\}ds,\tag{7}$$

上式中的积分是波赫纳(Bochner)积分. 关于波赫纳积分见附录5. 可以证明: 当 f、$v\in C^1([0,\infty);H)$时, (7)是(4)的解. 对于 f、v: $[0, T]\to H$ 为波赫纳可积时, (7)的右端关于 t 是连续的, 我

们称它为(4)的广义解，或简称为(4)的解．还可以证明：对于 $t\geqslant 0$，$\|G(t)\|\leqslant 1$.

一般地，考虑发展方程

$$\begin{aligned}&\frac{dy}{dt}+Ay=Bv, \quad 0<t,\\&y(0)=y_0,\end{aligned} \tag{8}$$

这里 $-A$ 是 H 上强连续算子半群 $G(\cdot)$ 的母元，$B\in\mathscr{L}(U, H)$，而 U 是希尔伯特空间.

$\mathscr{U}_{ad}$ 是 $L^2([0, T]; U)$ 中的凸闭集.

对于 $v(\cdot)\in\mathscr{U}_{ad}$，(8)的解是

$$y(t)=G(t)y_0+\int_0^t G(t-s)Bv(s)ds. \tag{9}$$

设 Q、$Q_1\in\mathscr{L}(H)$ 是自共轭的，$R\in\mathscr{L}(U)$ 也是自共轭的，并且存在 $\gamma>0$ 使得

$$\langle Rv, v\rangle_U\geqslant\gamma\|v\|_U^2, \quad Q\geqslant 0, \quad Q_1\geqslant 0. \tag{10}$$

对于 $v(\cdot)\in\mathscr{U}_{ad}$，相应于方程(8)考虑泛函

$$\begin{aligned}J(v)=&\langle Q_1y(T), y(T)\rangle_H+\int_0^T\{\langle Qy(t), y(t)\rangle_H\\&+\langle Rv(t), v(t)\rangle_U\}dt.\end{aligned} \tag{11}$$

根据表达式(11)和(9)，知道 $J(v)$ 是 $L^2([0, T]; U)$ 上的二次泛函．根据(10)，成立着

$$J(v)\geqslant\int_0^T\langle Rv(t), v(t)\rangle_U\, dt\geqslant\gamma\int_0^T\|v(t)\|_U^2 dt.$$

因此，根据定理 2-1，存在唯一的 $u(\cdot)\in\mathscr{U}_{ad}$，使得

$$J(u)\leqslant J(v) \tag{12}$$

对于 $v\in\mathscr{U}_{ad}$ 成立.

根据定理 2-2，成立着变分不等式

$$\langle Q_1 y(T, u), y(T, v) - y(T, u)\rangle_H$$
$$+\int_0^T \langle Ru(t), v(t) - u(t)\rangle_U dt$$
$$+\int_0^T \langle Qy(t, u), y(t, v) - y(t, u)\rangle_H dt \geqslant 0,$$

由上式和(9)得到

$$\langle Q_1 y(T, u), \int_0^T G(T-t)B(v(t)-u(t))dt\rangle_H$$
$$+\int_0^T \langle Qy(\sigma, u), \int_0^\sigma G(\sigma-t)B(v(t)-u(t))dt\rangle_H d\sigma$$
$$+\int_0^T \langle Ru(t), v(t)-u(t)\rangle_U dt \geqslant 0.$$

所以

$$\int_0^T \langle B^* G^*(T-t)Q_1 y(T, u), v(t)-u(t)\rangle_U dt$$
$$+\int_0^T \langle \int_t^T B^* G^*(\sigma-t)Qy(\sigma, u)d\sigma, v(t)-u(t)\rangle_U dt$$
$$+\int_0^T \langle Ru(t), v(t)-u(t)\rangle_U dt \geqslant 0.$$

记

$$\psi(t) = G^*(T-t)Q_1 y(T, u) + \int_t^T G^*(\sigma-t)Qy(\sigma, u)d\sigma, \tag{13}$$

那末上式成为

$$\int_0^T \langle B^*\psi(t) + Ru(t), v(t)-u(t)\rangle_U dt \geqslant 0, \quad \forall v(\cdot) \in \mathscr{U}_{ad}. \tag{14}$$

而 $\psi(\cdot)$ 可以视为方程

$$\frac{d\psi}{dt} - A^*\psi = -Qy(t, u), \quad 0 \leqslant t \leqslant T, \qquad \psi(T) = Q_1 y(T, u) \tag{15}$$

的解，这里 A^* 是 A 的共轭算子.

特别，当 $\mathscr{U}_{ad} = L^2([0, T]; U)$ 时，(14)变成

$$\int_0^T \langle B^*\psi(t)+Ru(t),\, v(t)\rangle_U\, dt=0, \quad \forall v(\cdot)\in L^2([0,\ T];U),$$

所以

$$B^*\psi(t)+Ru(t)=0$$

在$[0,\ T]$上几乎处处成立，即最优控制$u(\cdot)$适合

$$u(t)=-R^{-1}B^*\psi(t). \tag{16}$$

因此

$$y(t,\ u)=G(t)y_0-\int_0^t G(t-s)BR^{-1}B^*\psi(s)ds. \tag{17}$$

从而得证

定理 1 设$u(\cdot)$、$y(\cdot)=y(\cdot,\ u)$是最优控制问题(9)、(11)和(12)的解，那末不等式(14)对于$v(\cdot)\in\mathscr{U}_{ad}$成立，其中$\psi$由(13)决定。如果$\mathscr{U}_{ad}=L^2([0,\ T];U)$，那末最优控制由(16)决定。 ∎

进而在$\mathscr{U}_{ad}=L^2([0,\ T];U)$讨论反馈控制。

定理 2 设$P(\cdot):[0,\ T]\to\mathscr{L}(H)$满足积分方程

$$P(t)+\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)G(\sigma-t)d\sigma$$

$$=G^*(T-t)Q_1G(T-t)+\int_t^T G^*(\sigma-t)QG(\sigma-t)d\sigma, \tag{18}$$

那末
$$\langle P(0)y_0,\ y_0\rangle=\inf_{v\in\mathscr{U}_{ad}} J(v).$$

证 设$v(\cdot)\in\mathscr{U}_{ad}=L^2([0,\ T];U)$，而

$$y(t)=G(t)y_0+\int_0^t G(t-s)Bv(s)ds, \quad 0\leqslant t.$$

设$t\leqslant\sigma$，那末

$$G(\sigma-t)y(t)=G(\sigma)y_0+\int_0^t G(\sigma-s)Bv(s)ds$$

$$=G(\sigma)y_0+\int_0^\sigma G(\sigma-s)Bv(s)\,ds-\int_t^\sigma G(\sigma-s)Bv(s)ds,$$

因而成立着

$$y(\sigma)=G(\sigma-t)y(t)+\int_t^\sigma G(\sigma-s)Bv(s)ds. \tag{19}$$

(18)作用于 $y(t)$，并注意到(19)，得到

$$\begin{aligned}
&P(t)y(t)+\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma) \\
&\quad\cdot\left\{y(\sigma)-\int_t^\sigma G(\sigma-s)Bv(s)ds\right\}d\sigma \\
&=G^*(T-t)Q_1\left(y(T)-\int_t^T G(T-s)Bv(s)ds\right) \\
&\quad+\int_t^T G^*(\sigma-t)Q\left(y(\sigma)-\int_t^\sigma G(\sigma-s)Bv(s)ds\right)d\sigma.
\end{aligned}$$

所以

$$\begin{aligned}
P(t)y(t)&=G^*(T-t)Q_1y(T)+\int_t^T G^*(\sigma-t)Qy(\sigma)d\sigma \\
&\quad-\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)y(\sigma)d\sigma \\
&\quad+\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)\left(\int_t^\sigma G(\sigma-s)Bv(s)ds\right)d\sigma \\
&\quad-\int_t^T G^*(T-t)Q_1G(T-s)Bv(s)ds \\
&\quad-\int_t^T G^*(\sigma-t)Q\int_t^\sigma G(\sigma-s)Bv(s)ds\,d\sigma.
\end{aligned}$$

但是，上式右端后三项等于

$$\begin{aligned}
&\int_t^T G^*(s-t)\left\{\int_s^T G^*(\sigma-s)P(\sigma)BR^{-1}B^*P(\sigma)G(\sigma-s)d\sigma\right. \\
&\quad\left.-\int_s^T G^*(\sigma-s)QG(\sigma-s)d\sigma-G^*(T-s)Q_1G(T-s)\right\}Bv(s)ds \\
&=-\int_t^T G^*(s-t)P(s)Bv(s)ds,
\end{aligned}$$

所以

$$P(t)y(t)=G^*(T-t)Q_1y(T)+\int_t^T G^*(\sigma-t)Qy(\sigma)d\sigma$$

$$-\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)y(\sigma)d\sigma$$

$$-\int_t^T G^*(\sigma-t)P(\sigma)Bv(\sigma)d\sigma. \qquad (20)$$

因此

$$\langle P(t)y(t),\ y(t)\rangle=\langle Q_1y(T),\ G(T-t)y(t)\rangle$$

$$+\int_t^T\langle Qy(\sigma),\ G(\sigma-t)y(t)\rangle d\sigma$$

$$-\int_t^T\langle P(\sigma)BR^{-1}B^*P(\sigma)y(\sigma),\ G(\sigma-t)y(t)\rangle d\sigma$$

$$-\int_t^T\langle P(\sigma)Bv(\sigma),\ G(\sigma-t)y(t)\rangle d\sigma.$$

注意到(19)，由上式得到

$$\langle P(t)y(t),\ y(t)\rangle=\langle Q_1y(T),\ y(T)\rangle+\int_t^T\langle Qy(\sigma),\ y(\sigma)\rangle d\sigma$$

$$-\int_t^T\langle R^{-1}B^*P(\sigma)y(\sigma),\ B^*P(\sigma)y(\sigma)\rangle d\sigma$$

$$-\int_t^T\langle Q_1y(T),\ G(T-s)Bv(s)\rangle ds$$

$$-\int_t^T\left\langle Qy(\sigma),\ \int_t^\sigma G(\sigma-s)Bv(s)ds\right\rangle d\sigma$$

$$+\int_t^T\left\langle P(\sigma)BR^{-1}B^*P(\sigma)y(\sigma),\ \int_t^\sigma G(\sigma-s)Bv(s)ds\right\rangle d\sigma$$

$$-\int_t^T\langle P(\sigma)Bv(\sigma),\ y(\sigma)\rangle d\sigma$$

$$+\int_t^T\left\langle P(\sigma)Bv(\sigma),\ \int_t^\sigma G(\sigma-s)Bv(s)ds\right\rangle d\sigma.$$

根据(20)，上式右端第四、五、六、八项的和为

$$-\int_t^T\langle P(s)y(s),\ Bv(s)\rangle ds,$$

所以

$$\langle P(t)y(t),\ y(t)\rangle = \langle Q_1 y(T),\ y(T)\rangle + \int_t^T \langle Qy(\sigma),\ y(\sigma)\rangle d\sigma$$

$$- \int_t^T \langle R^{-1}B^*P(\sigma)y(\sigma),\ B^*P(\sigma)y(\sigma)\rangle d\sigma$$

$$- \int_t^T \langle P(s)y(s),\ Bv(s)\rangle ds$$

$$- \int_t^T \langle P(\sigma)Bv(\sigma),\ y(\sigma)\rangle d\sigma.$$

从而

$$\langle P(t)y(t),\ y(t)\rangle$$

$$= \langle Q_1 y(T),\ y(T)\rangle$$

$$+ \int_t^T \{\langle Qy(\sigma),\ y(\sigma)\rangle + \langle Rv(\sigma),\ v(\sigma)\rangle\} d\sigma$$

$$- \int_t^T \langle R(v(\sigma) + R^{-1}B^*P(\sigma)y(\sigma)),$$

$$v(\sigma) + R^{-1}B^*P(\sigma)y(\sigma)\rangle d\sigma.$$

因此，我们有

$$J(v(\cdot)) = \langle P(0)y_0,\ y_0\rangle$$

$$
\left.\begin{aligned}
y(t) &= G(t)y_0 - \int_0^t G(t-s)BR^{-1}B^*\psi(s)ds, \\
\psi(t) &= G^*(T-t)Q_1 y(T) + \int_t^T G^*(\sigma-t)Qy(\sigma)d\sigma,
\end{aligned}\right\} \quad 0\leqslant t\leqslant T, \tag{22}
$$

并且 $P(\cdot)$ 满足(18)，那末

$$
\psi(t) = P(t)y(t).
$$

证 由(22)得到：当 $t\leqslant\sigma$ 时

$$
y(\sigma) = G(\sigma-t)y(t) - \int_t^\sigma G(\sigma-s)BR^{-1}B^*\psi(s)ds,
$$

所以

$$
\begin{aligned}
\psi(t) = {} & G^*(T-t)Q_1G(T-t)y(t) \\
& - G^*(T-t)Q_1\int_t^T G(T-s)BR^{-1}B^*\psi(s)ds \\
& + \int_t^T G^*(\sigma-t)QG(\sigma-t)y(t)d\sigma \\
& - \int_t^T G^*(\sigma-t)Q\int_t^\sigma G(\sigma-s)BR^{-1}B^*\psi(s)ds\,d\sigma.
\end{aligned}
$$

从而根据(18)得到

$$
\begin{aligned}
\psi(t) = {} & P(t)y(t) \\
& + \int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)G(\sigma-t)d\sigma\, y(t) \\
& - \int_t^T G^*(s-t)\Big\{P(s) \\
& + \int_s^T G^*(\sigma-s)P(\sigma)BR^{-1}B^*P(\sigma)G(\sigma-s)d\sigma\Big\}BR^{-1}B^*\psi(s)ds \\
= {} & P(t)y(t) - \int_t^T G^*(s-t)P(s)BR^{-1}B^*\psi(s)ds \\
& + \int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)G(\sigma-t)y(t)d\sigma
\end{aligned}
$$

$$-\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)$$

$$\cdot\left(\int_t^\sigma G(\sigma-s)BR^{-1}B^*\psi(s)\,ds\right)d\sigma$$

$$=P(t)y(t)-\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*\psi(\sigma)\,d\sigma$$

$$+\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*P(\sigma)y(\sigma)\,d\sigma.$$

因此，成立着

$$\psi(t)-P(t)y(t)$$

$$=-\int_t^T G^*(\sigma-t)P(\sigma)BR^{-1}B^*(\psi(\sigma)-P(\sigma)y(\sigma))\,d\sigma.$$

由此得到不等式

$$\|\psi(t)-P(t)y(t)\|$$

$$\leqslant\int_t^T \|G^*(\sigma-t)P(\sigma)BR^{-1}B^*\|\cdot\|\psi(\sigma)-P(\sigma)y(\sigma)\|\,d\sigma.$$

根据格隆瓦尔不等式，由上式得到

$$\psi(t)-P(t)y(t)=0. \qquad \blacksquare$$

对于系统(8)，考虑泛函

$$J(v(\cdot))=\int_0^\infty\{\langle Qy(t),\ y(t)\rangle+\langle Rv(t),\ v(t)\rangle\}dt, \tag{23}$$

这里假设 $\mathscr{U}_{ad}=L^2([0,\ \infty);\ U)$，并且存在 $\alpha>0$ 和 M 使得

$$\|G(t)\|\leqslant Me^{-\alpha t},\qquad \forall 0\leqslant t<+\infty.$$

根据§2，存在 $u(\cdot)\in\mathscr{U}_{ad}$ 使得

$$J(u(\cdot))=\inf_{v(\cdot)\in\mathscr{U}_{ad}} J(v(\cdot)), \tag{24}$$

并且最优控制 $u(\cdot)\in\mathscr{U}_{ad}$ 由下述变分不等式

$$\int_0^\infty\{\langle Qy(t,\ u),\ y(t,\ v)-y(t,\ u)\rangle+\langle Ru(t),\ v(t)-u(t)\rangle\}dt$$

$$\geqslant 0,$$

$$\forall v(\cdot)\in\mathscr{U}_{ad}$$

表征．从而

$$\int_0^{\infty}\left\langle Qy(t,\ u),\ \int_0^t G(t-s)B(v(s)-u(s))ds\right\rangle dt$$

$$+\int_0^{\infty}\langle Ru(t),\ v(t)-u(t)\rangle dt\geqslant 0,$$

所以

$$\int_0^{\infty}\left\langle\int_s^{\infty} B^*G^*(\sigma-s)Qy(\sigma,\ u)d\sigma+Ru(s),\ v(s)-u(s)\right\rangle ds\geqslant 0.$$

由于 $\mathscr{U}_{ad}=L^2([0,\ +\infty);U)$，由上式得到

$$Ru(t)+\int_t^{\infty} B^*G^*(\sigma-t)Qy(\sigma,\ u)d\sigma=0. \tag{25}$$

置

$$\psi(t)=\int_t^{\infty} G^*(\sigma-t)Qy(\sigma,\ u)d\sigma, \tag{26}$$

那末(25)成为

$$Ru(t)+B^*\psi(t)=0,$$

从而

$$u(t)=-R^{-1}B^*\psi(t). \tag{27}$$

定理 4 设 $P^*=P\in\mathscr{L}(H)$ 适合

$$P=\int_0^{\infty} G^*(s)QG(s)ds-\int_0^{\infty} G^*(s)PBR^{-1}B^*PG(s)ds, \tag{28}$$

那末 $\qquad \langle Py_0,\ y_0\rangle=\inf\{J(v)\mid v(\cdot)\in\mathscr{U}_{ad}\},$

并且最优控制问题的解 $u(\cdot)$ 和 $y(\cdot,\ u)$ 适合

$$u(t)=-R^{-1}B^*Py(t,\ u). \tag{29}$$

证 由(28)知

$$P=\int_t^{\infty} G^*(\sigma-t)QG(\sigma-t)d\sigma$$

$$-\int_t^{\infty} G^*(\sigma-t)PBR^{-1}B^*PG(\sigma-t)d\sigma.$$

运用定理 2 中的方法，可以得到

$$\langle Py(t), y(t)\rangle = \int_t^\infty \{\langle Qy(\sigma), y(\sigma)\rangle + \langle Rv(\sigma), v(\sigma)\rangle\} d\sigma$$

$$- \int_t^\infty \langle R(v(\sigma) + R^{-1}B^*Py(\sigma)), v(\sigma) + R^{-1}B^*Py(\sigma)\rangle d\sigma. \tag{30}$$

所以

$$J(v(\cdot)) = \langle Py_0, y_0\rangle + \int_0^\infty \langle R(v(t) + R^{-1}B^*Py(t)), v(t) + R^{-1}B^*Py(t)\rangle dt \geqslant \langle Py_0, y_0\rangle.$$

因此 $\langle Py_0, y_0\rangle = \inf\{J(v(\cdot)) \mid v(\cdot) \in \mathscr{U}_{ad}\}$,

并且当 $u(\cdot)$和 $y(\cdot, u)$适合关系(29)时

$$J(u(\cdot)) = \langle Py_0, y_0\rangle. \qquad \blacksquare$$

习　题

1 具体推导关系(30).

2 试讨论热传导方程

$$\frac{\partial y}{\partial t} = \Delta y + v(t)g(x), \quad (t, x) \in (0, T) \times \Omega,$$

$$y(t, x) = 0, \qquad (t, x) \in \Sigma,$$

$$y(0, x) = y_0(x), \quad x \in \Omega$$

关于二次泛函

$$J(v(\cdot)) = \int_\Omega (y(T, x) - z_d(x))^2 dx + \int_0^T v^2(t) dt$$

取最小的最优控制问题，这里 $g(\cdot) \in L^2(\Omega)$, $z_d(\cdot) \in L^2(\Omega)$ 给定，$\mathscr{U}_{ad}$ 是 $L^2([0, T])$的凸闭集.

§5　时间最优控制

设 X 是巴拿赫空间，$G(\cdot)$是 X 上的强连续算子半群，即

$$G(t) \in \mathscr{L}(X), \qquad \forall t \geqslant 0,$$

$$G(t+s) = G(t)G(s), \quad \forall t, s \geqslant 0,$$

$$G(0)=I,$$

$$\lim_{t\to 0^+} G(t)x=x, \quad \forall x\in X.$$

例如，当 $A\in\mathscr{L}(X)$时，定义

$$G(t)=\sum_{j=0}^{\infty}\frac{(At)^j}{j!}, \tag{1}$$

它是 X 上的连续算子半群．(1)通常记为 $G(t)=e^{At}$.

对于 X 上的强连续算子半群 $G(\cdot)$，存在常数 M 和 ω 使得

$$\|G(t)\|\leqslant Me^{\omega t}, \quad \forall t\geqslant 0.$$

可以证明：存在 X 中的稠密集 $D(A)$，使得当 $x\in D(A)$时，极限

$$\lim_{t\to 0^+}\frac{G(t)x-x}{t}$$

存在，记它为 Ax．$D(A)$上的线性算子 A，称为 $G(\cdot)$的母元．可以证明，线性算子 A 是闭算子，即如果 $x_k\in D(A)$，并且

$$\lim_{k\to\infty} x_k=x, \quad \lim_{k\to\infty} Ax_k=y$$

存在，那末 $x\in D(A)$，并且 $y=Ax$.

关于算子半群，希尔、吉田耕作进行了详细讨论，有兴趣的读者请查阅有关专著．

法托里尼(H. O. Fattorini)讨论了 X 上的发展方程

$$\begin{gathered}\frac{dx}{dt}=Ax+v(t),\\ x(0)=x_0, \quad x(t_1)=x_1\end{gathered} \tag{2}$$

的时间最优控制问题，这里 x_0, $x_1\in X$ 给定，A 是算子半群 $G(\cdot)$ 的母元，$v(\cdot)$：$[0, t_1]\to X$ 是波赫纳(Bochner)可积的，并且 $\|v(t)\|\leqslant 1$，今后记它为 $v(\cdot)\in U_{ad}$.

$L^\infty([0, \infty); X)$表示 $v(\cdot)$：$[0, \infty)\to X$ 是有界的，并且它在任何有界区间上是波赫纳可积的．

对于 $v(\cdot)\in L^\infty([0, \infty); X)$，(2)的广义解是

$$x(t, v)=G(t)x_0+\int_0^t G(t-s)v(s)ds. \tag{3}$$

时间最优控制问题 选取 $u(\cdot)\in U_{ad}$，使得

$$x(t^*, u)=x_1,$$

并且如果 $x(t_1, v)=x_1$, $v(\cdot)\in U_{ad}$, 那末 $t^*\leqslant t_1$.

置

$$K(t)=\left\{x \mid x=\int_0^t G(t-s)v(s)ds,\ v(\cdot)\in L^\infty([0, +\infty); X)\right\}.$$

定理 1 当 $0<t<t'$ 时

$$K(t)=K(t').$$

证 因为

$$\int_0^t G(t-s)v(s)\,ds=\int_0^t G(\sigma)v(t-\sigma)d\sigma,$$

易知
$$K(t)\subset K(t').$$

又因为

$$\int_0^{t'} G(\sigma)v(\sigma)\,d\sigma=\int_0^t G(\sigma)v(\sigma)\,d\sigma+\int_0^{t'-t} G(t+s)v(t+s)ds$$
$$=\int_0^t G(\sigma)\left\{v(\sigma)+\frac{1}{t}G(t-\sigma)\int_0^{t'-t} G(s)v(t+s)ds\right\}d\sigma,$$

所以 $K(t')\subset K(t)$. 因此，$K(t')=K(t)$. ∎

记
$$K=K(t), \quad \forall t>0.$$

设 E 是 $[0, +\infty)$ 中的有界可测集，记

$$K_E(t)$$
$$=\left\{x \mid x=\int_0^t G(t-s)v(s)\chi_E(s)ds,\ v(\cdot)\in L^\infty([0, \infty); X)\right\}.$$

显然成立着 $K_E(t)\subset K(t)$, $\forall t>0$.

现在证明

定理 2 设 $t\in E$，使得

$$\lim_{h\to 0^+}\frac{1}{h}\int_{t-h}^t \chi_E(s)ds=1, \tag{4}$$

那末 $$K_E(t)=K.$$

证 由(4)得到，存在 $\delta>0$，使得当 $0<h'<h<\delta$ 时

$$\int_{t-h}^{t-h'}\chi_E(s)ds\geqslant\frac{1}{2}(h-h'),$$

取 $t_1=t-\frac{1}{2}\delta$, $t_2=(t+t_1)/2$, $\cdots$, $t_{n+1}=(t+t_n)/2$, $\cdots$, 那末

$$t_{n+1}-t_n=\frac{1}{2}(t_n-t_{n-1})\geqslant 0,$$

$$t-t_{n+1}=\frac{1}{2}(t-t_n)\geqslant 0,$$

$$\lim_{n\to\infty}t_n=t,$$

因此 $$\int_{t_n}^{t_{n+1}}\chi_E(s)ds\geqslant\frac{1}{2}(t_{n+1}-t_n).$$

如果 $x\in K$，那末存在 $v(\cdot)\in L^\infty([0,\ \infty);\ X)$ 使得

$$x=\int_{t_1}^{t}G(t-s)v(s)\,ds.$$

但是

$$\int_{t_1}^{t}G(t-s)v(s)\,ds=\sum_{n=1}^{\infty}\int_{t_n}^{t_{n+1}}G(t-s)v(s)\,ds$$

$$=\frac{\displaystyle\sum_{n=1}^{\infty}\int_{t_{n+1}}^{t_{n+2}}G(t-\sigma)\left(\chi_E(\sigma)\int_{t_n}^{t_{n+1}}G(\sigma-s)v(s)ds\right)d\sigma}{\displaystyle\int_{t_{n+1}}^{t_{n+2}}\chi_E(\sigma)d\sigma.}\tag{5}$$

置 $$\tilde{v}(\sigma)=\int_{t_n}^{t_{n+1}}G(\sigma-s)v(s)ds\Big/\int_{t_{n+1}}^{t_{n+2}}\chi_E(\sigma)d\sigma,$$

$$\forall\sigma\in[t_{n+1},\ t_{n+2}],$$

那末因为

$$\|\tilde{v}(\sigma)\|\leqslant C(t_{n+1}-t_n)\Big/\frac{1}{2}(t_{n+2}-t_{n+1})=4C,$$

这里 C 使得 $\|G(\sigma-s)v(s)\|\leqslant C$, $\forall 0\leqslant s\leqslant\sigma\leqslant t$, 所以

$$\tilde{v}(\cdot)\in L^\infty([0,\ +\infty);\ X).$$

从而由(5)得到

$$x=\int_{t_1}^{t} G(t-s)v(s)\,ds=\sum_{n=1}^{\infty}\int_{t_{n+1}}^{t_n} G(t-\sigma)\tilde{v}(\sigma)\chi_E(\sigma)d\sigma$$

$$=\int_{t_2}^{t} G(t-\sigma)\tilde{v}(\sigma)\chi_E(\sigma)d\sigma\in K_E(t).$$

因此得证 $K\subset K_E(t)$. ▌

在实变数函数论中证得, E 中使 (4) 不成立的点的全体是零集. 因此,对几乎所有的 $t\in E$,成立着

$$K_E(t)=K.$$

引理 1 如果 $x(t_1,\ v)=x_1$, $v(\cdot)\in U_{ad}$, 并且存在 $\varepsilon>0$ 使得 $\|v(t)\|\leqslant 1-\varepsilon$ 在 $[0,\ t_1]$ 上几乎处处成立, 那末 $v(\cdot)$ 不是最优的.

证 根据假设,成立着

$$x_1=G(t_1)x_0+\int_0^{t_1} G(t_1-s)v(s)ds$$

$$=G(t_1)x_0+\int_0^{t_1} G(\sigma)v(t_1-\sigma)d\sigma.$$

对于 $\tau<t_1$, 成立着

$$G(t_1)x_0-G(\tau)x_0=\frac{1}{\tau}\int_0^{\tau} G(\sigma)\{G(t_1-\sigma)x_0-G(\tau-\sigma)x_0\}d\sigma,$$

$$\int_{\tau}^{t_1} G(\sigma)v(t_1-\sigma)d\sigma=\int_0^{t_1-\tau} G(\tau+s)v(t_1-\tau-s)ds$$

$$=\frac{1}{\tau}\int_0^{\tau} G(\sigma)\left(\int_0^{t_1-\tau} G(\tau+s-\sigma)v(t_1-\tau-s)ds\right)d\sigma.$$

因此

$$x_1=G(\tau)x_0+\int_0^{\tau} G(\sigma)\Big\{v(t_1-\sigma)$$

$$+\frac{1}{\tau}(G(t_1-\sigma)x_0-G(\tau-\sigma)x_0)$$

$$+\frac{1}{\tau}\int_0^{t_1-\tau} G(\tau+s-\sigma)v(t_1-\tau-s)ds\Big\}d\sigma.$$

当 $t_1-\tau>0$ 充分小时，成立着

$$\frac{1}{\tau}\int_0^{t_1-\tau}\|G(\tau+s-\sigma)v(t_1-\tau+s)\|ds$$

$$\leqslant\frac{1}{\tau}\int_0^{t_1-\tau}\|G(\tau+s-\sigma)\|ds\leqslant\frac{\varepsilon}{2},$$

$$\frac{1}{\tau}\|G(t_1-\sigma)x_0-G(\tau-\sigma)x_0\|$$

$$\leqslant\frac{1}{\tau}\|G(\tau-\sigma)\|\|G(t_1-\tau)x_0-x_0\|<\frac{\varepsilon}{2}.$$

置

$$\tilde{v}(\sigma)=v(t_1-\sigma)+\frac{1}{\tau}(G(t_1-\sigma)x_0-G(\tau-\sigma)x_0)$$

$$+\frac{1}{\tau}\int_0^{t_1-\tau}G(\tau+s-\sigma)v(t_1-\tau-s)ds,$$

那末 $\tilde{v}(\cdot)$ 是波赫纳可积的，并且

$$\|\tilde{v}(\sigma)\|\leqslant 1-\varepsilon+\varepsilon=1,$$

$$x_1=G(\tau)x_0+\int_0^{\tau}G(\sigma)\tilde{v}(\sigma)d\sigma,$$

所以 x_1 可以用时间 τ 达到，即 $v(\cdot)$ 不是最优的. ∎

定理 3 如果 $u(\cdot)$ 和 $x(\cdot)=x(\cdot,u)$ 是时间最优控制问题的解，t^* 是最优时间，那末

$$\|u(t)\|=1 \tag{6}$$

在 $[0,t^*]$ 上几乎处处成立.

证 否则，存在 $\varepsilon>0$ 和可测集 E, $\mu(E)>0$，使得当 $t\in E$ 时

$$\|u(t)\|\leqslant 1-\varepsilon.$$

根据定理 2，存在 $s\in E$, $s<t^*$，使得 $K_E(s)=K=K(s)$，所以存在 $\tilde{v}(\cdot)\in L^\infty([0,\infty);X)$ 使得

$$\int_0^s G(s-\sigma)\tilde{v}(\sigma)\chi_E(\sigma)d\sigma=\int_0^s G(s-\sigma)u(\sigma)d\sigma.$$

置 $\quad v(t)=(1-\delta)u(t)+\tilde{v}(t)\chi_E(t)\delta,\quad \forall t\in[0,t^*].$

那末当 $t\in E$ 时

$$\|v(t)\|\leqslant(1-\delta)(1-\varepsilon)+\delta\|\tilde{v}(t)\|,$$

当 $t\bar{\in} E$ 时

$$\|v(t)\|\leqslant 1-\delta. \tag{7}$$

从而存在 $\varepsilon_1>0$，使得当 δ 充分小时，$\|v(t)\|\leqslant 1-\varepsilon_1$ 在 $[0, t^*]$ 上成立.

但是

$$\begin{aligned}x(s, v)&=G(s)x_0+\int_0^s G(s-\sigma)v(\sigma)d\sigma\\&=G(s)x_0+(1-\delta)\int_0^s G(s-\sigma)u(\sigma)d\sigma\\&\quad+\delta\int_0^s G(s-\sigma)\tilde{v}(\sigma)\chi_E(\sigma)d\sigma.\end{aligned}$$

上式右端最后一项等于

$$\delta\int_0^s G(s-\sigma)u(\sigma)d\sigma,$$

所以 $\quad x(s, v)=G(s)x_0+\int_0^s G(s-\sigma)u(\sigma)d\sigma=x(s, u).$

由于 $u(\cdot)$ 是把 x_0 迁移到 x_1 的最优控制，根据最优性原理，它是把 x_0 迁移到 $x(s, u)$ 的最优控制. 从而由 $x(s, v)=x(s, u)$ 知 $v(\cdot)$ 也是把 x_0 迁移到 $x(s, u)$ 的最优控制.

但是，根据(7)和引理1知，$v(\cdot)$ 不能是把 x_0 迁移到 $x(s, u)$ 的最优控制. 从而得到矛盾. 所以等式(6)在 $[0, t^*]$ 上几乎处处成立. ∎

一般地，讨论发展方程

$$\frac{dx}{dt}=Ax+b(t, v(t)), \tag{8}$$

$$x(0)=x_0,\ v(t)\in U,$$

这里 U 是巴拿赫空间 Z 中的紧集，称为控制区域；$b(\cdot, \cdot): \mathbb{R}\times Z\to X$ 是连续的；A 是强连续算子半群 $G(\cdot)$ 的母元.

如果 $v(\cdot):[0, t_1]\to U$ 是波赫纳可积的，就称 $v(\cdot)$ 是容许控制，记为 $v(\cdot)\in U_{ad}$.

当 $b(t, v)\equiv v$, $U=\{x\,|\,\|x\|\leqslant 1\}$ 时，(8) 成为 (2).

当 $v(\cdot)\in U_{ad}$ 时，我们称

$$x(t, v)=G(t)x_0+\int_0^t G(t-s)b(s, v(s))ds \tag{9}$$

为 (8) 的广义解，或简称为 (8) 的解.

记系统 (8) 的等时区域 $R(t)$ 为

$$R(t)=\{x\,|\,x=x(t, v),\ v(\cdot)\in U_{ad}\}.$$

根据波赫纳积分的性质，运用第二章类似的方法，可以证明

定理 4 当 $t\geqslant 0$ 时，$R(t)$ 的闭包 $\bar{R}(t)$ 是 X 中的凸集.

证 设 $x_j\in\bar{R}(t)$ $(j=1, 2)$, $\alpha\in(0, 1)$.

对于 $\varepsilon>0$，存在 $v_j(\cdot)\in U_{ad}$，使得

$$\|x_j-x(t, v_j)\|<\frac{\varepsilon}{2},$$

所以

$$\|\alpha x_1+(1-\alpha)x_2-\alpha x(t, v_1)-(1-\alpha)x(t, v_2)\|<\frac{\varepsilon}{2}.$$

但是

$$\begin{aligned}&\alpha x(t, v_1)+(1-\alpha)x(t, v_2)\\&\quad=G(t)x_0+\int_0^t G(t-s)b(s, v_2(s))ds\\&\qquad+\alpha\int_0^t G(t-s)\{b(s, v_1(s))-b(s, v_2(s))\}ds.\end{aligned}$$

根据 $\alpha\in(0, 1)$，存在 $E_\alpha^s\subset[0, t]$，使得

$$\begin{aligned}&\alpha\int_0^t G(t-s)\{b(s, v_1(s))-b(s, v_2(s))\}ds\\&\quad=\int_{E_\alpha^s} G(t-s)\{b(s, v_1(s))-b(s, v_2(s))\}ds+\eta,\end{aligned}$$

$$|\eta| < \frac{\varepsilon}{2}.$$

所以

$$\left\| \alpha x_1 + (1-\alpha)x_2 - G(t)x_0 - \int_{E_\alpha^\varepsilon} G(t-s)b(s, v_1(s))ds \right.$$

$$\left. - \int_{[0,t]\setminus E_\alpha^\varepsilon} G(t-s)b(s, v_2(s))ds \right\| < \varepsilon.$$

置
$$v^\varepsilon(s) = \begin{cases} v_1(s), & \forall s \in E_\alpha^\varepsilon, \\ v_2(s), & \forall s \in [0, t]\setminus E_\alpha^\varepsilon, \end{cases}$$

那末 $v^\varepsilon(\cdot) \in U_{ad}$, 并且

$$\left\| \alpha x_1 + (1-\alpha)x_2 - G(t)x_0 - \int_0^t G(t-s)b(s, v^\varepsilon(s))ds \right\| < \varepsilon.$$

所以 $\alpha x_1 + (1-\alpha)x_2 \in \bar{R}(t)$. ∎

设在 X 中给定了目标集 Q, 它是凸闭集, 并且含有内点. 如果讨论下述最优控制问题

$$t^* = \inf\{\tilde{t} \mid \bar{R}(\tilde{t}) \cap Q \neq \emptyset\}.$$

运用第二章 § 4 类似的方法, 可以得到

$$\bar{R}(t^*) \cap Q \neq \emptyset,$$

$$\bar{R}(t) \cap Q = \emptyset, \quad \forall 0 \leqslant t < t^*.$$

如果存在 $u(\cdot) \in U_{ad}$, 使得 $x(t^*, u) \in Q$, 我们就称 $u(\cdot)$ 和 $x(\cdot) = x(\cdot, u)$ 为把 x_0 迁移到 Q 的时间最优控制问题的解. 这时, 可以证明 $x(t^*, u)$ 是 Q 和 $\bar{R}(t^*)$ 的公共边界点.

根据巴拿赫空间凸集的分离性定理(因为 Q 含有内点), 存在 $\lambda \in X^*$, $\|\lambda\| = 1$, 使得

$$\langle \lambda, x - x(t^*) \rangle \geqslant 0, \quad \forall x \in Q, \tag{10}$$

$$\langle \lambda, x(t^*, v) - x(t^*) \rangle \leqslant 0, \quad \forall v(\cdot) \in U_{ad}.$$

由上式得到

$$\int_0^{t^*} \langle \lambda, G(t^*-t)\{b(t, v(t)) - b(t, u(t))\} \rangle dt \leqslant 0,$$

从而

$$\int_0^{t^*}\langle G^*(t^*-t)\lambda,\ b(t,\ v(t))-b(t,\ u(t))\rangle dt\leqslant 0,$$

$$\forall v(\cdot)\in U_{ad}. \tag{11}$$

置

$$\psi(t)=G^*(t^*-t)\lambda,$$

那末

$$\frac{d\psi(t)}{dt}=-A^*\psi(t),\quad \psi(t^*)=\lambda, \tag{12}$$

而(11)成为

$$\int_0^{t^*}\langle\psi(t),\ b(t,\ v(t))-b(t,\ u(t))\rangle dt\leqslant 0,$$

$$\forall v(\cdot)\in U_{ad}, \tag{13}$$

如果 t 是 $\langle\psi(\cdot),\ b(\cdot,\ u(\cdot))\rangle$ 的勒贝格点，$h>0$，$v\in U$，置

$$v(s)=\begin{cases}v, & \forall s\in[t-h,\ t+h],\\ u(s), & \forall s\in[0,\ t^*]\setminus[t-h,\ t+h],\end{cases}$$

那末 $v(\cdot)\in U_{ad}$，由(13)得到

$$\int_{t-h}^{t+h}\langle\psi(s),\ b(s,\ v)-b(s,\ u(s))\rangle ds\leqslant 0,$$

两端除以 $2h$，并令 $h\to 0$，得到

$$\langle\psi(t),\ b(t,\ v)-b(t,\ u(t))\rangle\leqslant 0,$$

所以

$$\max_{v\in U}\langle\psi(t),\ b(t,\ v)\rangle=\langle\psi(t),\ b(t,\ u(t))\rangle. \tag{14}$$

这是最大值条件. 而不等式(10)是

$$\langle\psi(t^*),\ x-x(t^*)\rangle\geqslant 0\qquad \forall x\in Q, \tag{15}$$

它是横截条件.

当 $b(t,\ v)=Bv$，$B\in\mathscr{L}(Z,\ X)$，U 是 Z 中的凸闭集时，可以

证明: $R(t)$是 X 中的凸闭集.

这时，当 Q 为 X 中的凸闭集且含有内点时，最优控制适合最大值条件.

关于系统(8)时间最优控制问题的讨论，详见李训经、姚允龙的论文〈3〉.

附录1 巴拿赫空间

一 线性空间

设对于集 X 定义了元素之间的加法及元素与数的乘法，即对于 x 和 $y \in X$，存在唯一的 $x+y \in X$ 与之对应，且满足

$$x+y=y+x,$$

存在 0 元满足

$$x+0=x,$$

对于 $x \in X$，存在唯一的 $-x$ 与之对应，使得

$$x+(-x)=0;$$

对于 $x \in X$ 和 $\alpha \in \mathbb{R}$，存在唯一的 $\alpha x \in X$ 与之对应，且满足

$$\alpha(x+y)=\alpha x+\alpha y, \qquad \forall x, y \in X, \alpha \in \mathbb{R};$$

$$(\alpha+\beta)x=\alpha x+\beta x, \qquad \forall \alpha, \beta \in \mathbb{R}, x \in X,$$

$$\alpha(\beta x)=(\alpha\beta)x, \qquad \forall \alpha, \beta \in \mathbb{R}, x \in X,$$

$$1 \cdot x=x, \qquad \forall x \in X,$$

$$0 \cdot x=0, \qquad \forall x \in X,$$

应注意的是：上式右端 $0 \in X$，左端 $0 \in \mathbb{R}$.

这时，称 X 是线性空间.

例如，$[\alpha, \beta]$ 上勒贝格可积函数的全体，在通常的加法和数乘下构成一线性空间. 又如，$[\alpha, \beta]$ 上连续函数的全体，在通常的加法和数乘下也构成一线性空间.

二 线性赋范空间

如果 X 是线性空间，并且对于 $x \in X$ 有非负数 $\|x\|$ 与之对应，且满足

$$\|x\| \geqslant 0, \quad \|x\|=0 \text{ 的充要条件是 } x=0,$$

$$\|\alpha x\| = |\alpha| \|x\|, \quad \forall \alpha \in \mathbb{R}, \quad x \in X,$$

$$\|x+y\| \leqslant \|x\| + \|y\|, \quad \forall x, y \in X,$$

就称 X 是线性赋范空间.

当 $x_n \in X$, $n=0, 1, 2, \cdots$, $x \in X$ 满足

$$\lim_{n\to\infty} \|x_n - x\| = 0$$

时,称 x 是 $\{x_n\}$ 的极限. 如果 $\{x_n\}$ 有极限, 那末

$$\lim_{n, m\to\infty} \|x_n - x_m\| = 0. \tag{1}$$

但是, 一般说来, 由上式不能肯定 $\{x_n\}$ 有极限. 例如 X 是由有理数的全体构成的,不能由 $\lim\limits_{n, m\to\infty} |x_n - x_m| = 0$ 推出 $\{x_n\}$ 有极限.

如果线性赋范空间中任一满足(1)的序列 $\{x_n\}$ 都在 X 中有极限 x,就称 X 是巴拿赫(S. Banach)空间.

例如, L 是 $[\alpha, \beta]$ 上勒贝格可积函数 $x(\cdot)$ 的全体,在 L 中定义加法和数乘按通常意义,而 $x(\cdot)$ 的范数定义为

$$\|x\| = \int_\alpha^\beta |x(t)| dt.$$

那末 $\|x(\cdot)\|$ 满足范数的三个要求,并且当

$$\lim_{n, m\to\infty} \int_\alpha^\beta |x_n(t) - x_m(t)| dt = 0$$

时,存在 $x(\cdot) \in L$,使得

$$\lim_{n\to\infty} \int_\alpha^\beta |x_n(t) - x(t)| dt = 0,$$

所以, L 是一巴拿赫空间.

又例如 $L^p([\alpha, \beta])$ $(p>1)$. $[\alpha, \beta]$ 上勒贝格可积函数 $x(\cdot)$,且适合

$$\int_\alpha^\beta |x(t)|^p dt < +\infty,$$

的函数全体, 简记为 L^p. 在通常的加法和数乘下构成线性空间.这是因为

$$|x(t)+y(t)|^p\leqslant 2^p(|x(t)|^p+|y(t)|^p),$$

所以当 $x(\cdot)$、$y(\cdot)\in L^p([\alpha,\beta])$时，$x(\cdot)+y(\cdot)\in L^p([\alpha,\beta])$. 在 $L^p([\alpha,\beta])$引进范数

$$\|x(\cdot)\|_p=\left(\int_\alpha^\beta|x(t)|^p dt\right)^{\frac{1}{p}}.$$

那末可以证明三角不等式

$$\|x(\cdot)+y(\cdot)\|_p\leqslant\|x(\cdot)\|_p+\|y(\cdot)\|_p.$$

事实上，如果α，$\beta\geqslant 0$，$\dfrac{1}{p}+\dfrac{1}{q}=1$，那末

$$\alpha\beta\leqslant\frac{\alpha^p}{p}+\frac{\beta^q}{q}.$$

从而由 $u(\cdot)\in L$，$v(\cdot)\in L^p$，推知不等式

$$\frac{|u(t)||v(t)|}{\|u\|_p\cdot\|v\|_q}\leqslant\frac{1}{p}\frac{|u(t)|^p}{\|u\|_p^p}+\frac{1}{q}\frac{|v(t)|^q}{\|v\|_q^q}$$

右端在$[\alpha,\beta]$上是可积的，且积分等于

$$\frac{1}{p}+\frac{1}{q}=1,$$

所以，上式左端是可积的，从而

$$\int_\alpha^\beta|u(t)v(t)|dt\leqslant\|u\|_p\cdot\|v\|_q$$

$$=\left(\int_\alpha^\beta|u(t)|^p dt\right)^{\frac{1}{p}}\cdot\left(\int_\alpha^\beta|v(t)|^q dt\right)^{\frac{1}{q}}.$$

该不等式称为霍尔窦(Hölder)不等式.

如果 $x(\cdot)$、$y(\cdot)\in L^p$，那末$|x(\cdot)|+|y(\cdot)|\in L^p$，从而

$$(|x(\cdot)|+|y(\cdot)|)^{p-1}\in L^q,$$

而

$$\begin{aligned}|x(t)+y(t)|^p&\leqslant(|x(t)|+|y(t)|)^p\\&=|x(t)|\cdot(|x(t)|+|y(t)|)^{p-1}\\&\quad+|y(t)|(|x(t)|+|y(t)|)^{p-1}.\end{aligned}$$

对上式右端每一项，应用霍尔窦不等式，得到

$$\int_\alpha^\beta |x(t)+y(t)|^p dt \leqslant \int_\alpha^\beta (|x(t)|+|y(t)|)^p dt$$
$$\leqslant \|x(\cdot)\|_p \cdot \left(\int_\alpha^\beta (|x(t)|+|y(t)|)^p dt\right)^{\frac{1}{q}}$$
$$+\|y(\cdot)\|_p \left(\int_\alpha^\beta (|x(t)|+|y(t)|)^p dt\right)^{\frac{1}{q}}.$$

从而 $\left\{\int_\alpha^\beta (|x(t)|+|y(t)|)^p dt\right\}^{\frac{1}{p}} \leqslant \|x(\cdot)\|_p + \|y(\cdot)\|_p.$

因此三角不等式

$$\|x(\cdot)+y(\cdot)\|_p \leqslant \|x(\cdot)\|_p + \|y(\cdot)\|_p$$

成立．所以 $L^p([\alpha, \beta])\,(p>1)$ 也是线性赋范空间．也可以证明 $L^p([\alpha,\beta])$ 是完备的．从而 $L^p([\alpha, \beta])\,(p\geqslant 1)$ 是巴拿赫空间．

再如空间 $C([\alpha, \beta])$．$[\alpha, \beta]$ 上的连续函数 $x(\cdot)$ 的全体，以通常的加法和数乘构成线性空间，而 $x(\cdot)$ 赋以范数

$$\|x(\cdot)\|_C = \max\{|x(t)|\,|\,t\in[\alpha, \beta]\}.$$

它构成完备的线性赋范空间．

三 连续线性泛函

如果 $f: X\to\mathbb{R}$ 是线性的，即满足

$$f(x+y)=f(x)+f(y), \quad \forall x, y\in X,$$
$$f(\alpha x)=\alpha f(x), \quad \forall \alpha\in\mathbb{R},\ x\in X,$$

就称 f 是 X 上的线性泛函．如果存在常数 $M>0$，使得

$$|f(x)|\leqslant M\|x\|,$$

就称 f 是 X 上的连续线性泛函．

设 $p>1$，$\frac{1}{p}+\frac{1}{q}=1$．如果 $x(\cdot):[\alpha, \beta]\to\mathbb{R}^n$ 是可积的，并且 $\|x(\cdot)\|\in L^p([\alpha, \beta])$，就称 $x(\cdot)\in L^p([\alpha, \beta]; \mathbb{R}^n)$．在通常的加法和数乘下，$L([\alpha, \beta]; \mathbb{R}^n)$ 是线性空间．若赋以范数

$$\|x(\cdot)\|_p = \left(\int_\alpha^\beta \|x(t)\|^p dt\right)^{\frac{1}{p}},$$

从而，$L^p([\alpha, \beta]; \mathbb{R}^n)$是巴拿赫空间．对于$y(\cdot)\in L^q([\alpha, \beta]; \mathbb{R}^n)$，由积分

$$f(x)=\int_\alpha^\beta \langle x(t), y(t)\rangle dt \tag{2}$$

定义了$L^p([\alpha, \beta]; \mathbb{R}^n)$上的线性泛函．这里$\langle\cdot, \cdot\rangle$是$\mathbb{R}^n$中的内积．利用霍尔窦不等式得到

$$|f(x)|\leqslant\int_\alpha^\beta \|x(t)\|\cdot\|y(t)\|dt\leqslant\|x(\cdot)\|_p\cdot\|y(\cdot)\|_q,$$

知f是$L^p([\alpha, \beta]; \mathbb{R}^n)$上的连续线性泛函．

黎斯(F. Riesz)证明：$L^q([\alpha, \beta])$ $(p>1)$的连续线性泛函f，必存在$y(\cdot)\in L^q([\alpha, \beta])$，使得

$$f(x)=\int_\alpha^\beta x(t)y(t)dt.$$

而$L^p([\alpha, \beta]; \mathbb{R}^n)$ $(p>1)$上的连续线性泛函必可表示为(2)．证明见夏道行等《实变数函数论与泛函分析概要》[3]．

线性赋范空间X上的连续线性泛函的全体X^*，在通常的加法和数乘下构成线性空间，如果规定X上的连续线性泛函$f: X\to\mathbb{R}$的范数为

$$\|f\|=\sup\frac{|f(x)|}{\|x\|},$$

那末X^*是线性赋范空间．可以证明：X^*是完备的，即X^*是巴拿赫空间．称X^*是线性赋范空间X的共轭空间．

事实上，如果

$$\lim_{n, m\to\infty}\|f_n-f_m\|=0,$$

那末由

$$|f_n(x)-f_m(x)|=|(f_n-f_m)(x)|\leqslant\|f_n-f_m\|\|x\|$$

推知$\{f_n(x)\}$(当$x\in X$固定时)是有极限的，我们定义

$$f(x)=\lim_{n\to\infty}f_n(x),$$

容易证明 $f: X \to \mathbb{R}$ 是线性泛函. 并且由

$$|f_n(x)| \leqslant \|f_n\| \cdot \|x\| \leqslant M\|x\|$$

推知

$$|f(x)| \leqslant M\|x\|,$$

这里 $M>0$ 满足 $\|f_n\| \leqslant M$.

应当注意的是: 在证明 X^* 的完备性时, 并没有用到 X 的完备性.

X^* 的共轭空间记为 X^{**}. 对于 $x \in X$, $f \in X^*$, 记

$$\langle f, x\rangle = f(x).$$

那末 $\langle \cdot, x\rangle: X^* \to \mathbb{R}$ 是线性的, 并且

$$|\langle f, x\rangle| \leqslant \|f\| \cdot \|x\|,$$

即

$$\frac{|\langle f, x\rangle|}{\|f\|} \leqslant \|x\|.$$

因此, $\langle \cdot, x\rangle$ 是 X^* 上的连续线性泛函, 它是 X^{**} 的元素. 由此, 空间 X 的每一元 x 对应于 X^{**} 的某一元. 这就是说, 可以把 X 嵌入到 X^{**}. 如果 $X^{**} = X$, 就称 X 是自反的.

根据 $L^p([\alpha, \beta]; \mathbb{R}^n)$ $(p>1)$ 的连续线性泛函的黎斯表现定理知, 它的共轭空间为 $L^q([\alpha, \beta]; \mathbb{R}^n)$, 其中 q 适合

$$\frac{1}{p} + \frac{1}{q} = 1.$$

而 $L^q([\alpha, \beta]; \mathbb{R}^n)$ 的共轭空间为 $L^p([\alpha, \beta]; \mathbb{R}^n)$. 因此, $L^p([\alpha, \beta]; \mathbb{R}^n)$ 是自反的巴拿赫空间.

可以证明: $L([\alpha, \beta]; \mathbb{R}^n)$ 的共轭空间是 $L^\infty([\alpha, \beta]; \mathbb{R}^n)$ (区间 $[\alpha, \beta]$ 上取值于 $\mathbb{R}^n$ 的有界可积函数的全体). 但是, $L^\infty([\alpha, \beta]; \mathbb{R}^n)$ 的共轭空间不是 $L([\alpha, \beta]; \mathbb{R}^n)$. 所以, $L([\alpha, \beta]; \mathbb{R}^n)$ 不是自反的巴拿赫空间.

四　弱收敛与弱*收敛

如果 x_n、$x \in X$, 使得对任何 $f \in X^*$, 成立着

$$\lim_{n \to \infty} f(x_n) = f(x),$$

就称$\{x_n\}$在 X 上弱收敛于 x. 显然, 如果

$$\lim_{n\to\infty}\|x_n-x\|=0,$$

即$\{x_n\}$强收敛于 x, 那末$\{x_n\}$弱收敛于 x. 但反过来不一定对. 例如$[0, 2\pi]$上的平方可积函数全体$L^2([0, 2\pi])$, 根据勒贝格定理, 对于 $x(\cdot)\in L^2([0, 2\pi])$, 成立着

$$\lim_{n\to\infty}\int_0^{2\pi}x(t)\cos(nt)dt=0,$$

但是

$$\int_0^{2\pi}\cos^2(nt)dt=\pi$$

不趋于 0, 即 $\{\cos(nt)\}$ 在 $L^2([0, 2\pi])$上弱收敛于 0, 但不强收敛于 0.

对于 $f_n\in X^*$, 如果对任何 $x\in X$, 成立着

$$\lim_{n\to\infty}f_n(x)=f(x),$$

就称 f_n 弱*收敛于 f.

例 如果 $x_n(\cdot)\in L^\infty([\alpha, \beta])$, $x(\cdot)\in L^\infty([\alpha, \beta])$适合

$$\lim_{n\to\infty}\int_\alpha^\beta x_n(t)y(t)dt=\int_\alpha^\beta x(t)y(t)dt$$

对于 $y(\cdot)\in L^1([\alpha, \beta])$皆成立, 就称 $x_n(\cdot)$ 在 L^∞ 上弱*收敛于 $x(\cdot)$.

巴拿赫空间 X 中的集 E 称为是闭的, 如果由

$$x_n\in E,\quad \lim_{n\to\infty}\|x_n-x\|=0,$$

能推出 $x\in E$; 集 E 称为是弱闭的, 如果由 $x_n\in E$, $n=0, 1, 2, \cdots$, x_n 弱收敛于 x, 能得到 $x\in E$.

显然, 弱闭集必是闭集. 但反之不一定对. 例如, $L^2([0, 2\pi])$中的单位球面 B:

$$\|x(\cdot)\|_2=\int_0^{2\pi}|x(t)|^2dt=1.$$

显然, 如果 $x_n(\cdot)\in B$, $\|x_n(\cdot)-x(\cdot)\|_2\to 0$, 那末由

$$\int_0^{2\pi}|x_n(t)|^2dt\to\int_0^{2\pi}|x(t)|^2dt$$

知 $x(\cdot)\in B$，所以 B 是 $L^2([0,\ 2\pi])$ 中的闭集．但是

$$\frac{1}{\pi}\sin(n\cdot)$$

的范数为 1，它的弱极限 0 不在 B 中，即 B 不是弱闭的．

定理 7 设 X 是可分的巴拿赫空间，$f_n\in X^*$，并且存在 $M>0$ 使得

$$\|f_n\|\leqslant M,$$

那末 $\{f_n\}$ 存在弱*收敛的子列．

注 称巴拿赫空间 X 是可分的，如果存在 $x_k\in X$，$k=1,2,\cdots$，使得对于 $x\in X$ 和任何 $\delta>0$，存在 x_k，使得

$$\|x-x_k\|<\delta.$$

即集 $\{x_k\}$ 的极限点全体与 X 一致．

证 由于 $\{f_n(x_1)\}$ 是有界数列，所以存在收敛子列，记为 $\{f_{1,m}(x_1)\}$，即 $\{f_{1,m}(x_1)\}\subset\{f_n(x_1)\}$，并且

$$\lim_{m\to\infty}f_{1,m}(x_1)$$

存在．再由于 $\{f_{1,m}(x_2)\}$ 是有界数列，存在子列 $\{f_{2,m}(x_2)\}\subset\{f_{1,m}(x_2)\}$，使得极限

$$\lim_{m\to\infty}f_{2,m}(x_2)$$

存在，这时显然极限

$$\lim_{m\to\infty}f_{2,m}(x_1)$$

也存在．逐次进行下去，得子列 $\{f_{k,m}\}_{m=1}^{\infty}$，使得极限

$$\lim_{m\to\infty}f_{k,m}(x_i)$$

对于 $i=1,\ 2,\ \cdots,\ k$ 皆存在．

取 $f^{(k)}=f_{k,k}$，因为除有限项外

$$\{f_{k,k}\}_{k=1}^{\infty}$$

是 $\{f_{l,m}\}_{m=1}^{\infty}$ 的子列，所以对于 x_l，皆有极限

$$\lim_{k\to\infty} f^{(k)}(x_l)$$

存在.

现在证明：对于 $x\in X$，极限 $\lim\limits_{k\to\infty} f^{(k)}(x)$ 是存在的，这只要证明

$$\lim_{n,m\to\infty} |f^{(n)}(x)-f^{(m)}(x)|=0.$$

事实上，由于

$$|f^{(n)}(x)-f^{(m)}(x)|\leqslant|f^{(n)}(x)-f^{(n)}(x_l)|$$
$$+|f^{(n)}(x_l)-f^{(m)}(x_l)|+|f^{(m)}(x_l)-f^{(m)}(x)|,$$

对 $x\in X$ 和 $\varepsilon>0$，取 x_l 满足

$$\|x-x_l\|<\varepsilon,$$

那末 $$|f^{(n)}(x)-f^{(n)}(x_l)|\leqslant M\varepsilon,$$
$$|f^{(m)}(x)-f^{(m)}(x_l)|\leqslant M\varepsilon,$$

从而 $$|f^{(n)}(x)-f^{(m)}(x)|\leqslant 2M\varepsilon+|f^{(n)}(x_l)-f^{(m)}(x_l)|,$$

所以 $$\overline{\lim_{n,m\to\infty}}|f^{(n)}(x)-f^{(m)}(x)|\leqslant 2M\varepsilon,$$

但 $\varepsilon>0$ 是任意的，所以

$$\lim_{n,m\to\infty}|f_{n,n}(x)-f_{m,m}(x)|=0,$$

即极限 $$\lim_{n\to\infty} f^{(n)}(x)=\lim_{n\to\infty} f_{n,n}(x)$$

存在．设 $$f(x)=\lim_{n\to\infty} f_{n,n}(x),$$

那末 $f: X\to\mathbb{R}$ 是线性的，并且由于

$$|f_{n,n}(x)|\leqslant M\|x\|,$$

得到 $$|f(x)|\leqslant M\|x\|.$$

即 f： $X\to\mathbb{R}$ 是连续线性泛函，它是 $\{f_n\}$ 的子列 $\{f_{n,n}\}$ 的弱*极限. ∎

定理1表明：X^* 中的有界集是弱*致密的，即能在其中选取弱*收敛的子列.

把定理1应用到 $L^\infty([\alpha, \beta]; \mathbb{R}^n)$ 就得到：有界可积函数中的有界集是弱*致密的. 我们在证明等时区域的闭性时曾利用该结论.

同样，$L^p([\alpha, \beta]; \mathbb{R}^n)$ $(p>1)$ 中的有界集是弱致密的. 这是因为：$L^p([\alpha, \beta]; \mathbb{R}^n)$ 是 $L^q([\alpha, \beta]; \mathbb{R}^n)$ $(q=p/(p-1))$ 的共轭空间，$L^q([\alpha, \beta]; \mathbb{R}^n)$ 是可分的，并且 $L^q([\alpha, \beta]; \mathbb{R}^n)$ 是自反的. 所以 $L^p([\alpha, \beta]; \mathbb{R}^n)$ 的弱*收敛和弱收敛是一致的.

定理2 设 X 是可分的巴拿赫空间，那末存在 $f_n \in X^*$，$\|f_n\| \leqslant 1$，使得对任何 $f_0 \in X^*$，$\|f_0\| \leqslant 1$，在 $\{f_n\}$ 中有子列 $\{f_{n_j}\}$ 使得 f_{n_j} 弱*收敛于 f_0.

证 因为 X 是可分的，存在 $\{x_n\}$ 在 X 中是稠密的（即它的闭包是 X）.

考虑 $\varphi_n: X^* \to \mathbb{R}^n$ 为

$$\varphi_n(f) = (f(x_1), f(x_2), \cdots, f(x_n)).$$

因为 $\mathbb{R}^n$ 是可分的，存在 $f_{n,k} \in X^*$，使得 $\|f_{n,k}\| \leqslant 1$，而

$$\{\varphi_n(f_{n,k})\}_{k=1}^{\infty}$$

在 $\{\varphi_n(f) \mid f \in X^*, \|f\| \leqslant 1\}$ 中是稠密的.

这时 $\{f_{n,k}\}$ $(k=1, 2, \cdots; n=1, 2, \cdots)$ 是可数的，$f_{n,k} \in X^*$，$\|f_{n,k}\| \leqslant 1$.

对于 $f_0 \in X^*$，$\|f_0\| \leqslant 1$，在 $\{f_{n,k}\}_{k=1}^{\infty}$ 中取 f_{n,m_n}，使得

$$|f_{n,m_n}(x_j) - f_0(x_j)| < \frac{1}{n}, \quad j=1, 2, \cdots, n.$$

这时，对于任一 x_j，成立着

$$\lim_{n\to\infty} f_{n,m_n}(x_j) = f_0(x_j).$$

和定理1一样，因为 $\{x_j\}$ 在 X 中稠密，所以

$$\lim_{n\to\infty} f_{n,m_n}(x) = f_0(x).$$

即 f_{n,m_n} 弱*收敛于 f_0. ∎

定理 2 表明：当 X 可分时，X^* 的单位球 $\{f \mid \|f\| \leqslant 1\}$ 也是可分的.

五 凸集

设 E 是线性空间 X 中的子集，如果对于 $x\in E$, $y\in E$ 和 $\lambda\in[0, 1]$, 成立着

$$\lambda x + (1-\lambda)y \in E,$$

就称 E 是 X 中的凸集.

当 X 是巴拿赫空间时，如果存在 $x_0\in E$ 和 $\delta>0$, 使得

$$\{x \mid \|x-x_0\| < \delta\} \subset E,$$

就称 E 含有内点，并且称 x_0 是 E 的内点. 我们今后主要讨论巴拿赫空间 X 中含有内点的凸集.

例如，X 中半径为 $r>0$ 的球 $\{x \mid \|x\| \leqslant r\}$ 是含有内点的凸集.

不失一般性，可以设 0 是凸集 E 的内点. 这时，对于 $x\in X$, 存在数 $\lambda>0$, 使得 x/λ 是 E 中的点，定义

$$p(x) = \inf\{\lambda>0 \mid x/\lambda \in E\},$$

那末 $p(\cdot): X \to \mathbb{R}$ 具有下述性质：

1° $p(x) \geqslant 0$, $\forall x\in X$,

2° $p(\alpha x) = \alpha p(x)$, $\forall x\in X$, $\alpha \geqslant 0$,

3° $p(x+y) \leqslant p(x)+p(y)$, $\forall x, y\in X$.

性质 1° 是显然的；由 $0\in E$, 知 $p(0)=0$, 所以 2° 对 $\alpha=0$ 成立，对于 $\alpha>0$, 由于

$$p(\alpha x) = \inf\left\{\lambda>0 \,\middle|\, \frac{1}{\lambda}(\alpha x) \in E\right\} = \inf\left\{\lambda>0 \,\middle|\, \frac{\alpha}{\lambda}x \in E\right\}$$

$$= \alpha \inf\left\{\mu>0 \,\middle|\, \frac{1}{\mu}x \in E\right\} = \alpha p(x),$$

所以性质 2° 成立. 关于性质 3° 的证明如下. 如果 $\lambda>0$, $\mu>0$ 使得 $\frac{1}{\lambda}x\in E$, $\frac{1}{\mu}y\in E$, 那末根据 E 的凸性得到

$$\frac{1}{\lambda+\mu}(x+y)=\frac{\lambda}{\lambda+\mu}\cdot\frac{x}{\lambda}+\frac{\mu}{\lambda+\mu}\cdot\frac{y}{\mu}\in E,$$

因此

$$\begin{aligned}p(x+y)&=\inf\left\{r>0\,\middle|\,\frac{1}{r}(x+y)\in E\right\}\\&\leqslant\inf\left\{\lambda+\mu\,\middle|\,\lambda>0,\ \mu>0,\ \frac{1}{\lambda}x\in E,\ \frac{1}{\mu}y\in E\right\}\\&=\inf\left\{\lambda>0\,\middle|\,\frac{1}{\lambda}x\in E\right\}+\inf\left\{\mu>0\,\middle|\,\frac{1}{\mu}y\in E\right\}.\end{aligned}$$

所以性质 3° 成立.

$p(\cdot)$称为凸集 E 的闵可夫斯基(H. Minkowski)函数.

定理 3 设 E 是 X 中的凸集, 0 是 E 的内点, $p(\cdot)$是 E 的闵可夫斯基函数, 那末当且仅当 x 是 E 的内点时

$$p(x)<1,$$

当且仅当 x 是 $X\setminus E$ 的内点时

$$p(x)>1.$$

证 当 $x\in E$ 时, 显然成立 $p(x)\leqslant1$.

如果 x 是 E 的内点, 那末当 $\delta_1>0$ 充分小时, 点 $(1+\delta_1)x$ 属于 E, 从而

$$p(x)=\inf\left\{\lambda>0\,\middle|\,\frac{1}{\lambda}x\in E\right\}\leqslant\frac{1}{1+\delta_1}<1;$$

反之, 设 $p(x)<1$. 记 $2\delta_1=1-p(x)>0$, 这时

$$\frac{1}{1-\delta_1}x=\frac{1}{p(x)+\delta_1}x\in E,$$

由于 0 是 E 的内点, 所以存在 $\delta>0$, 使得当 $\|y\|<\delta$ 时 $y\in E$. 从而根据 E 的凸性得到

$$\delta_1 y+(1-\delta_1)\cdot\frac{1}{1-\delta_1}x\in E,$$

即 $x+\delta_1 y\in E$. 因此, 以 x 为心 $\delta_1\delta$ 为半径的球

$$\{\tilde{x}\mid\|\tilde{x}-x\|<\delta_1\delta\}\subset E,$$

即 x 是 E 的内点.

如果 x 是 $X\setminus E$ 的内点, 那末存在 $\delta_1>0$ 使得 $(1-\delta_1)x\bar{\in}E$. 这样, 如果 $\lambda>0$, $\frac{1}{\lambda}x\in E$, 就有 $\frac{1}{\lambda}<1-\delta_1$, 从而 $\lambda\geqslant\frac{1}{1-\delta_1}$, 因此 $p(x)\geqslant 1$. 反之, 设 $p(x)>1$. 记 $2\delta_1=p(x)-1>0$, 这时 $1+\delta_1=p(x)-\delta_1$, 所以

$$\frac{1}{1+\delta_1}x\bar{\in}E. \tag{3}$$

这时, 球 $\{\tilde{x}\mid\|\tilde{x}-x\|<\delta_1\delta\}\subset X\setminus E$. 否则, 存在 y 满足

$$\|y\|<\delta,\quad x+\delta_1 y\in E.$$

因为点 $(-y)\in E$, 从而

$$\frac{\delta_1}{1+\delta_1}(-y)+\frac{1}{1+\delta_1}(x+\delta_1 y)=\frac{1}{1+\delta_1}x\in E.$$

它与(3)相矛盾. 所以球 $\{\tilde{x}\mid\|\tilde{x}-x\|<\delta_1\delta\}$ 位在 $X\setminus E$ 中, 即 x 是 $X\setminus E$ 的内点. ∎

因此, 闵可夫斯基函数 $p(\cdot)$ 可以用来刻画 x 是否是 E 的内点.

定理 4 设 X_1 是线性空间 X 的子空间, f_1 是 X_1 上的线性泛函, 满足 $f_1(x)\leqslant p(x)$, $\forall x\in X_1$, 其中 $p(\cdot): X\to\mathbb{R}$ 满足

$$p(x+y)\leqslant p(x)+p(y),\quad \forall x, y\in X,$$
$$p(\alpha x)=\alpha p(x),\quad \forall x\in X,\quad \alpha\geqslant 0.$$

那末在 X 上必存在线性泛函 $f: X\to\mathbb{R}$, 满足

$$f(x)\leqslant p(x),\quad \forall x\in X,$$
$$f(x)=f_1(x),\quad \forall x\in X_1.$$

定理 4 称为汉恩(H. Hahn)-巴拿赫线性泛函的延拓定理. 证

明见《线性泛函分析入门》[5].

定理 5 设 E 是巴拿赫空间 X 的凸集，0 是 E 的内点，$p(\cdot)$ 是 E 的闵可夫斯基函数，x_0 不是 E 的内点，那末存在非零的连续线性泛函 $f: X\to\mathbb{R}$，满足

$$f(x_0)=p(x_0),$$
$$f(x)\leqslant p(x),\quad \forall x\in X. \tag{4}$$

证 设
$$X_1=\{\lambda x_0 \mid \lambda\in\mathbb{R}\},$$
那末 X_1 是 X 的线性子空间，在 X_1 上定义
$$f_1(\lambda x_0)=\lambda p(x_0),$$
那末 f_1 是 X_1 上的非零线性泛函，并且当 $\lambda\geqslant 0$ 时
$$f_1(\lambda x_0)=p(\lambda x_0), \tag{5}$$
当 $\lambda<0$ 时
$$f_1(\lambda x_0)=\lambda p(x_0)\leqslant 0\leqslant p(\lambda x_0).$$
因此，当 $x\in X_1$ 时
$$f_1(x)\leqslant p(x).$$

根据定理 4，存在线性泛函 $f: X\to\mathbb{R}$，满足
$$f(x)=f_1(x),\quad \forall x\in X_1,$$
$$f(x)\leqslant p(x),\quad \forall x\in X. \tag{6}$$
而由 (5) 和 (6) 得到
$$f(x_0)=p(x_0).$$

因为 0 是 E 的内点，所以存在 $\delta>0$，使得当 $\|x\|<\delta$ 时
$$x\in E,$$
从而 $p(x)\leqslant 1$. 而由 (4) 得到 $f(-x)\leqslant p(-x)$，所以
$$f(x)\geqslant -p(-x).$$
从而
$$-p(-x)\leqslant f(x)\leqslant p(x).$$
因此，当 $\|x\|<\delta$ 时 $|f(x)|\leqslant 1$，即 f 在 $x=0$ 处是连续的. ∎

定理 6 设 E 和 F 是巴拿赫空间 X 的凸集，E 含有内点，F 的点都不是 E 的内点，那末存在非零的连续线性泛函，使得

$$\sup\{f(x)\,|\,x\in E\}\leqslant\inf\{f(y)\,|\,y\in F\}.$$

即 f 把 E 和 F 分在两侧.

证　置 $\widetilde{E}=E-F$, 那末 $\widetilde{E}$ 是含有内点 $\widetilde{x}_0$ 的凸集, 0 不是 $\widetilde{E}$ 的内点. 从而集 $\widetilde{E}-\widetilde{x}_0$ 是含有 0 为内点的凸集, $-\widetilde{x}_0$ 不是它的内点, 根据定理 5, 存在非零的连续泛函 f, 使得

$$f(-\widetilde{x}_0)=p(-\widetilde{x}_0),$$

$$f(x)\leqslant p(x),\qquad \forall x\in X.$$

这里 $p(\cdot)$ 是 $\widetilde{E}-\widetilde{x}_0$ 的闵可夫斯基函数. 特别, 当 $\widetilde{x}\in\widetilde{E}-\widetilde{x}_0$ 时

$$f(\widetilde{x})\leqslant p(\widetilde{x})\leqslant 1\leqslant p(-\widetilde{x}_0)=f(-\widetilde{x}_0).$$

因此, 当 $x\in E$, $y\in F$ 时 $x-y-\widetilde{x}_0\in\widetilde{E}-\widetilde{x}_0$, 从而

$$f(x-y-\widetilde{x}_0)\leqslant f(-\widetilde{x}_0),$$

即
$$f(x)\leqslant f(y).$$ ∎

这是巴拿赫空间中的凸集分离性定理.

设 $f\in X^*$, $\|f\|=1$, $r\in\mathbb{R}$, 称 X 中满足

$$f(x)\leqslant r$$

的 x 的全体为半空间. 显然, 半空间是闭凸集. 定理 5 表明: E 和 F 分在两个半空间中.

由于凸集的交是凸集, 所以有限个或无限个半空间的交集是闭凸集.

反之, 如果 E 是 X 中的闭凸集, 那末 E 是有限个或无限个半空间的交集.

事实上, 对于 $x_0\overline{\in} E$, 存在 $\delta>0$ 使得球 $\{x\,|\,\|x-x_0\|<\delta\}$ 位在 $X\setminus E$ 中. 对 E 和球 $\{x\,|\,\|x-x_0\|<\delta\}$ 应用定理 6, 得到: 存在 $f\in X^*$, $\|f\|=1$, 使得

$$\begin{aligned}\sup\{f(x)\,|\,x\in E\}&\leqslant\inf\{f(\widetilde{x})\,|\,\|\widetilde{x}-x_0\|<\delta\}\\&=f(x_0)+\inf\{f(\widetilde{x}-x_0)\,|\,\|\widetilde{x}-x_0\|<\delta\},\end{aligned}$$

从而
$$\sup\{f(x)\,|\,x\in E\}\leqslant f(x_0)-\delta.$$

因此，E 位在闭的半空间

$$H_{x_0}=\{x|f(x)\leqslant f(x_0)-\delta\}$$

中，从而
$$E\subset\bigcap\{H_{x_0}|x_0\overline{\in}E\}.$$

但是，显然 $x_0\overline{\in}H_{x_0}$. 所以只要 $\tilde{x}\overline{\in}E$, 必有

$$\tilde{x}_0\overline{\in}\bigcap\{H_{x_0}|x_0\overline{\in}E\}.$$

即
$$E\supset\bigcap\{H_{x_0}|x_0\overline{\in}E\}.$$

因此
$$E=\bigcap\{H_{x_0}|x_0\overline{\in}E\}.$$ ▌

附录 2　变分学基础

本附录讨论古典变分学的必要条件.

首先证明下述引理. 它是由 DuBois-Reymond 于 1879 年公布的.

引理　设 $m(\cdot):[t_0, t_1]\to\mathbb{R}$ 是逐段连续的, 如果等式

$$\int_{t_0}^{t_1} m(t)\dot{\eta}(t)\,dt=0 \tag{1}$$

对于适合 $\eta(t_0)=\eta(t_1)=0$ 的逐段光滑的 $\eta(\cdot)$ 皆成立, 那末 $m(t)$ 在 $[t_0, t_1]$ 上是常值的.

证　记

$$c=\int_{t_0}^{t_1} m(t)\,dt/(t_1-t_0),$$

那末

$$\eta(t)=\int_{t_0}^{t}(m(\tau)-c)\,d\tau \tag{2}$$

满足引理的条件, 并且

$$\int_{t_0}^{t_1} c\dot{\eta}(t)\,dt=c(\eta(t_1)-\eta(t_0))=0.$$

由 (1) 得到

$$\int_{t_0}^{t_1}(m(t)-c)\dot{\eta}(t)\,dt=0,$$

把 (2) 代入上式得

$$\int_{t_0}^{t_1}(m(t)-c)^2\,dt=0.$$

从而对于 $m(\cdot)$ 的连续点 t, 成立着

$$m(t)-c=0.$$

因此, 对于 $m(\cdot)$ 的可能的不连续点, 上式也成立. ▮

现讨论古典变分学的最简单问题.

设 $f:[t_0, t_1]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 是连续可微的，它在点 (t, x, r) 处的值为 $f(t, x, r)$. 设 $x(\cdot):[t_0, t_1]\to\mathbb{R}$ 是逐段光滑的，即 $x(\cdot)$ 连续且除 (t_0, t_1) 中的有限个点外 $x(\cdot)$ 的导数是连续的，并且在 $\dot{x}(\cdot)$ 不连续点 $\bar{t}$ 处，极限 $\dot{x}(\bar{t}-0)$ 和 $\dot{x}(\bar{t}+0)$ 是存在的. 如果 $x(t_0)=x_0$, $x(t_1)=x_1$，就称曲线 $x=x(\cdot)$ 是连结给定点 (t_0, x_0) 和 (t_1, x_1) 的.

对于连结 (t_0, x_0) 和 (t_1, x_1) 的曲线 $x=x(\cdot)$，考虑泛函

$$J(x(\cdot))=\int_{t_0}^{t_1} f\left(t, x(t), \frac{dx(t)}{dt}\right)dt. \qquad (3)$$

显然，泛函(3)是依赖于函数 $x(\cdot)$ 的. 考虑泛函的极值，是变分学的基本问题.

对于连结 (t_0, x_0) 和 (t_1, x_1) 的曲线 $x(\cdot)$ 和 $\bar{x}(\cdot)$，我们可以定义它们的零阶距离是

$$d_0[x(\cdot), \bar{x}(\cdot)]=\max_{t\in[t_0, t_1]}|x(t)-\bar{x}(t)|;$$

它们的1阶距离是

$$d_1[x(\cdot), \bar{x}(\cdot)]=d_0[x(\cdot), \bar{x}(\cdot)]+\sup_{t\in[t_0, t_1]\setminus D}|\dot{x}(t)-\dot{\bar{x}}(t)|,$$

这里 D 是 $\dot{x}(\cdot)$ 和 $\dot{\bar{x}}(\cdot)$ 的不连续点组成的集. 由此，可以定义 $x(\cdot)$ 的0阶 δ 邻域 $N_0(\delta, x(\cdot))$ 和1阶 δ 邻域 $N_1(\delta, x(\cdot))$. 显然，如果 $\bar{x}(\cdot)\in N_1(\delta, x(\cdot))$，那末 $\bar{x}(\cdot)\in N_0(\delta, x(\cdot))$. 但反过来是不对的.

如果存在 $\delta>0$，使得当 $\bar{x}(\cdot)\in N_1(\delta, x(\cdot))$ 时

$$J(x(\cdot))\leqslant J(\bar{x}(\cdot)), \qquad (4)$$

就称 $x(\cdot)$ 使泛函 $J(x(\cdot))$ 达到弱极小值；如果(4)对于 $\bar{x}(\cdot)\in N_0(\delta, x(\cdot))$ 成立，就称它使泛函(3)达到强极小值；如果(4)对于任何连结 (t_0, x_0) 和 (t_1, x_1) 的 $\bar{x}(\cdot)$ 皆成立，则称 $x(\cdot)$ 使(3)达到最

小值.

微积分学中求解函数极值的方法是建立在函数的微分在函数达到极值的点上恒等于零的基础上的，求解泛函的极值问题的方法也是类似的，为此，我们先引入泛函的变分概念，它类似于函数的微分的概念。

适合条件

$$L(\alpha x_1(\cdot)+\beta x_2(\cdot))=\alpha L(x_1(\cdot))+\beta L(x_2(\cdot))$$

的泛函 $L(x(\cdot))$称为线性泛函，此处 α 和 β 是任意常数。

定义 若泛函(3)在函数 $x(\cdot)$ 上的增量

$$\Delta J=J(\bar{x}(\cdot))-J(x(\cdot))$$

能表为下式 $$\Delta J=L(x(\cdot),\ \delta x)+o(d_0),$$

其中 $L(x(\cdot),\ \delta x)$是关于 $x(\cdot)$的增量 $\delta x=\bar{x}(\cdot)-x(\cdot)$(又称 δx 为 $x(\cdot)$的变分)的线性泛函，$d_0=d_0[x(\cdot),\ \bar{x}(\cdot)]$，则称 $L(x(\cdot),\ \delta x)$ 为泛函 $J(x(\cdot))$的在函数 $x(\cdot)$的变分，并记为 δJ.

由此可见

$$\Delta J=\delta J+o(d_0),$$

即泛函的变分 δJ 是泛函 $J(x(\cdot))$ 的改变量 ΔJ 关于 $x(\cdot)$的改变量(变分)的线性主部，这类似于微分学中函数的微分概念.

在微积分学中，函数 $f(x)$在点 x 上的微分可由式

$$df(x)=\left[\frac{\partial}{\partial\varepsilon}f(x+\varepsilon\Delta x)\right]_{\varepsilon=0}$$

计算. 类似地，由变分的定义可推得公式

$$\delta J(x(\cdot))=\left(\frac{\partial}{\partial\varepsilon}\Delta J\right)_{\varepsilon=0}=\left[\frac{\partial}{\partial\varepsilon}J(x(\cdot)+\varepsilon\delta x)\right]_{\varepsilon=0}.$$

现在推导使泛函(3)达到弱极小值的 $x(\cdot)$适合的必要条件.

设 $\eta(\cdot):[t_0,\ t_1]\to\mathbb{R}$ 是逐段光滑的，并且

$$\eta(t_0)=\eta(t_1)=0.$$

那末对于 $\delta>0$，存在 $\bar{\varepsilon}>0$ 使得当 $|\varepsilon|<\bar{\varepsilon}$ 时

$$\max_{t\in[t_0,t_1]}|\varepsilon\eta(t)|+\sup_{t\in[t_0,t_1]\setminus D}|\varepsilon\dot{\eta}(t)|<\delta,$$

这里 D 是 $\dot{\eta}(\cdot)$ 不连续点的集. 因此，当 $\varepsilon\in(-\bar{\varepsilon},\bar{\varepsilon})$ 时

$$\bar{x}(\cdot)=x(\cdot)+\varepsilon\eta(\cdot)\in N_1(\delta,x(\cdot)),$$

从而
$$J(x(\cdot))\leqslant J(x(\cdot)+\varepsilon\eta(\cdot)).$$

记
$$F(\varepsilon)=J(x(\cdot)+\varepsilon\eta(\cdot));$$

那末
$$F(0)\leqslant F(\varepsilon),$$

从而由于 $F(\cdot)$ 在 $\varepsilon=0$ 处达到极小值，根据微积分得到

$$F'(0)=0,\ \text{i.e.}\ \delta J(x(\cdot))=0,$$

也就是，泛函(3)在 $x(\cdot)$ 上达到弱极小值，则泛函在 $x(\cdot)$ 上的变分必恒等于零.

下面计算 $F'(0)$. 因为

$$F(\varepsilon)=\int_{t_0}^{t_1}f(t,\ x(t)+\varepsilon\eta(t),\ \dot{x}(t)+\varepsilon\dot{\eta}(t))dt,$$

所以

$$F'(0)=\int_{t_0}^{t_1}\left\{\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial x}\eta(t)+\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial r}\dot{\eta}(t)\right\}dt.$$

因此，如果 $x(\cdot)$ 使 J 达到弱极小，那末对于适合引理的 $\eta(\cdot)$，成立着

$$\int_{t_0}^{t_1}\left\{\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial x}\eta(t)+\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial r}\dot{\eta}(t)\right\}dt=0,\qquad(5)$$

由于

$$\int_{t_0}^{t_1}\frac{\partial f(t,\ x(t),\ \dot{x}(t))}{\partial x}\eta(t)dt=-\int_{t_0}^{t_1}\left(\int_{t_0}^{t}\frac{\partial f(\tau,\ x(\tau),\ \dot{x}(\tau))}{\partial x}d\tau\right)\dot{\eta}(t)dt,$$

所以由(5)得到

$$\int_{t_0}^{t_1}\left\{\frac{\partial f(t, x(t), \dot{x}(t))}{\partial r}-\int_{t_0}^{t}\frac{\partial f(\tau, x(\tau), \dot{x}(\tau))}{\partial x}d\tau\right\}\dot{\eta}(t)dt$$
$$=0.$$

根据引理和上式，得到

定理 1 设 $x(\cdot)$ 使泛函(3)达到弱极小值，那末

$$\frac{\partial f(t, x(t), \dot{x}(t))}{\partial r}-\int_{t_0}^{t}\frac{\partial f(\tau, x(\tau), \dot{x}(\tau))}{\partial x}d\tau\equiv c, \tag{6}$$

在 $[t_0, t_1]$ 上成立，其中 c 是某一常数. ∎

由(6)，对于 $\dot{x}(\cdot)$ 的连续点 $t\in[t_0, t_1]$，成立着

$$\frac{d}{dt}\left(\frac{\partial f(t, x(t), \dot{x}(t))}{\partial r}\right)-\frac{\partial f(t, x(t), \dot{x}(t))}{\partial x}=0. \tag{7}$$

如果使(3)取极小值的 $x(\cdot)$ 是光滑的，那末 $x(\cdot)$ 在区间 $[t_0, t_1]$ 上适合微分方程 (7). 微分方程 (7) 称为变分问题的欧拉方程. 而(6)是欧拉方程的积分形式.

定理 1 表明，泛函(3)取弱极值的函数 $x(\cdot)$ 适合欧拉方程. 有时称欧拉方程满足边值条件

$$x(t_0)=x_0, \qquad x(t_1)=x_1$$

的解为变分问题的极值曲线. 必须指出，我们只是证明了使泛函取弱极值的曲线是极值曲线. 但极值曲线并不一定使泛函(3)取弱极值.

根据强极值和弱极值的定义，如果 $x(\cdot)$ 使泛函(3)取强极值，那末它使(3)取弱极值，因此 $x(\cdot)$ 适合欧拉方程 (6).

Karl Weierstrass 对于使泛函(3)取强极小值的 $x(\cdot)$，进一步推导了 $x(\cdot)$ 所适合的条件. 他引进了函数

$$E:[t_0, t_1]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R},$$

E 在 (t, x, r, q) 处的值为

$$E(t,\ x,\ r,\ q)=f(t,\ x,\ q)-f(t,\ x,\ r)$$
$$-(q-r)f_r(t,\ x,\ r),$$

并证明了下述定理.

定理 2 如果 $x(\cdot)$ 使泛函 (3) 取强极小值, $\dot{x}(\cdot)$ 在 $t\in(t_0,\ t_1)$ 处是光滑的, 那末

$$E(t,\ x(t),\ \dot{x}(t),\ q)\geqslant 0 \tag{8}$$

对于 $q\in\mathbb{R}$ 皆成立. 如果 $\dot{x}(\cdot)$ 在 $t\in(t_0,\ t_1)$ 处是不连续的, 那末上式对于用 $\dot{x}(t\pm 0)$ 代替 $\dot{x}(t)$ 仍成立.

证 设 a 是 $x(\cdot)$ 的光滑点, 取 $b>a$ 使得 $x(\cdot)$ 在 $[a,\ b]$ 中是光滑的.

对于充分小的 $\varepsilon>0$, 定义 $x_\varepsilon(\cdot)$ 如下:

$$x_\varepsilon(t)=\begin{cases}x(t) & \forall t\in[t_0,\ a]\cup[b,\ t_1],\\ x(a)+q(t-a), & \forall t\in[a,\ a+\varepsilon],\\ \varphi(t,\ \varepsilon), & \forall t\in[a+\varepsilon,\ b],\end{cases}$$

这里

$$\varphi(t,\ \varepsilon)=x(t)+\frac{x(a)+\varepsilon q-x(a+\varepsilon)}{b-a-\varepsilon}(b-t). \tag{9}$$

这时 $x_\varepsilon(\cdot)$ 在 $[t_0,\ t_1]$ 上是逐段光滑的, 适合边值条件

$$x_\varepsilon(t_0)=x_0,\qquad x_\varepsilon(t_1)=x_1.$$

它与 $x(\cdot)$ 仅在区间 $[a,\ b]$ 上不同, 并且

$$\max_{t\in[t_0,t_1]}|x_\varepsilon(t)-x(t)|=\max_{t\in[a,b]}|x_\varepsilon(t)-x(t)|$$
$$\leqslant\varepsilon|q|+\max_{t\in[a,a+\varepsilon]}|x(a)-x(t)|.$$

所以, 当 $\varepsilon>0$ 充分小时 $x_\varepsilon(\cdot)\in N_0(\delta,\ x(\cdot))$.

又由 (9) 知道

$$\varphi(t,\ 0)=x(t),\qquad \forall t\in[a,\ b],$$

$$\varphi_\varepsilon(t,\ 0)=\frac{\partial\varphi(t,\ \varepsilon)}{\partial\varepsilon}\bigg|_{\varepsilon=0}=\lim_{\varepsilon\to 0}\frac{\varepsilon q-(x(a+\varepsilon)-x(a))}{\varepsilon(b-a-\varepsilon)}(b-t)$$

$$=\frac{q-\dot{x}(a)}{b-a}(b-t),\qquad \forall t\in(a,\ b].$$

由于 $x_\varepsilon(\cdot)\in N_0(\delta,\ x(\cdot))$，所以

$$\Phi(\varepsilon)=J(x_\varepsilon(\cdot))-J(x(\cdot))\geqslant 0,$$

并且

$$\Phi(0)=0.$$

从而

$$\Phi'(0)\geqslant 0. \tag{10}$$

但是

$$\Phi(\varepsilon)=\int_a^{a+\varepsilon}\{f(t,\ x(a)+q(t-a),\ q)-f(t,\ x(t),\ \dot{x}(t))\}dt$$

$$+\int_{a+\varepsilon}^b\{f(t,\ \varphi(t,\ \varepsilon),\ \dot{\varphi}(t,\ \varepsilon))-f(t,\ x(t),\ \dot{x}(t))\}dt.$$

所以

$$\Phi'(0)=\lim_{\varepsilon\to 0}\frac{\Phi(\varepsilon)}{\varepsilon}$$

$$=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^{a+\varepsilon}\{f(t,\ x(a)+q(t-a),\ q)-f(t,\ x(t),\dot{x}(t))\}dt$$

$$+\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_a^b\{f(t,\ \varphi(t,\ \varepsilon),\ \dot{\varphi}(t,\ \varepsilon))-f(t,\ x(t),\ \dot{x}(t))\}dt$$

$$+\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{a+\varepsilon}^a\{f(t,\ \varphi(t,\ \varepsilon),\ \dot{\varphi}(t,\ \varepsilon))-f(t,x(t),\dot{x}(t))\}dt.$$

上式右端第一项等于

$$f(a,\ x(a),\ q)-f(a,\ x(a),\ \dot{x}(a)),$$

第三项等于

$$-f(a,\ \varphi(a,\ 0),\ \dot{\varphi}(a,\ 0))+f(a,\ x(a),\ \dot{x}(a))=0,$$

第二项等于

$$\int_a^b \left\{ \frac{\partial f(t, x(t), \dot{x}(t))}{\partial x} \varphi_\varepsilon(t,0) + \frac{\partial f(t, x(t), \dot{x}(t))}{\partial r} \dot{\varphi}_\varepsilon(t,0) \right\} dt$$

$$= \int_a^b \left\{ \frac{\partial f(t, x(t), \dot{x}(t))}{\partial x} - \frac{d}{dt}\left(\frac{\partial f(t, x(t), \dot{x}(t))}{\partial r} \right) \right\} \varphi_\varepsilon(t,0)\, dt$$

$$+ \frac{\partial f(t, x(t), \dot{x}(t))}{\partial r} \varphi_\varepsilon(t, 0) \Big|_a^b$$

$$= -f_r(a, x(a), \dot{x}(a))(q - \dot{x}(a)).$$

因此

$$\Phi'(0) = f(a, x(a), q) - f(a, x(a), \dot{x}(a)) - (q - \dot{x}(a)) f_r(a, x(a), \dot{x}(a)) \geqslant 0.$$

对于 $\dot{x}(\cdot)$ 的不连续点 $\bar{t} \in (t_0, t_1)$，当 $\varepsilon > 0$ 充分小时，$\bar{t} \pm \varepsilon$ 是 $\dot{x}(\cdot)$ 的连续点，从而成立

$$E(\bar{t} \pm \varepsilon, x(\bar{t} \pm \varepsilon), \dot{x}(\bar{t} \pm \varepsilon), q) \geqslant 0.$$

令 $\varepsilon \to 0$，就得到

$$E(\bar{t}, x(\bar{t}), \dot{x}(\bar{t} \pm 0), q) \geqslant 0. \qquad \blacksquare$$

以上只是古典变分学中的简单问题的两个定理，对于更一般的变分学问题，相应地有本书正文中的定理 III-3-1 和定理 III-3-2。

附录 3 布劳维尔不动点定理

定理 1 设 $\mathbb{B}^n=\{x\in\mathbb{R}^n \mid \|x\|\leqslant 1\}$, $f:\mathbb{B}^n\to\mathbb{B}^n$ 是连续的, 那末存在 $x_0\in\mathbb{B}^n$ 满足

$$f(x_0)=x_0, \tag{1}$$

即 x_0 是 f 的不动点.

这是 L. E. J. Brouwer 于 1910 年在 Mathematics Annal 69 卷上公布的结果. 它的推广形式是

定理 2 设 Ω 是 $\mathbb{R}^n$ 中的有界凸闭集, $f:\Omega\to\Omega$ 是连续的, 那末存在 $x_0\in\Omega$ 满足(1), 即 f 存在不动点.

首先指出, $n=1$ 时定理 1 是平凡的. 事实上, 置

$$\varphi(x)=x-f(x),$$

那末 $\varphi(\cdot):[-1, 1]\to\mathbf{R}$ 是连续的, 并且

$$\varphi(-1)=-1-f(-1)\leqslant 0,$$
$$\varphi(1)=1-f(1)\geqslant 0.$$

所以存在 $x_0\in[-1, 1]$, 使得 $\varphi(x_0)=0$, 即 x_0 是 f 的不动点. 对于 $n>1$, Brouwer 不动点定理的证明有多种. 下面的证明具有构造性质, 即给出不动点的计算方法. 这是选自刊登于 Mathematics of Computation, Vol. 32(1978)上周修义、J. Mallet-Paret、J. A. Yorke 的论文的.

为引述周修义等的方法, 我们需要关于可微映照的沙得 (A. Sard) 定理.

设 f: $\mathbb{R}^n\to\mathbb{R}^p$ 是光滑的, $y\in\mathbb{R}^p$, 如果对于 $x\in f^{-1}(y)$ 皆有 $Df(x)$ 的值域

$$\text{Range}\, Df(x)=\mathbb{R}^p,$$

就称 y 是 f 的正则值. 当 y 不是 f 的正则值时，y 称为 f 的临界值，这时存在 $x\in f^{-1}(y)$ 使得

$$Df(x)$$

的秩小于 p. 这里 $Df(x)$ 是 f 在 x 处的一阶偏导数构成的 $n\times p$ 阶矩阵.

沙得定理 如果 f: $\mathbb{R}^n\to\mathbb{R}^p$ 具有 r 阶连续偏导数，$r>\max(0, n-p)$，那末对几乎所有的 $y\in\mathbb{R}^p$，y 是 f 的正则值，即 f 的临界值的全体在 $\mathbb{R}^p$ 中是零测度集.

这是有关可微映照的结果，我们不拟给出它的证明，请读者参阅有关整体分析、可微映照的专著.

现利用它来证明

定理 3 设 V 和 U 分别是 $\mathbb{R}^q$ 和 $\mathbb{R}^m$ 中的开集，映照 ϕ: $V\times U\to\mathbb{R}^p$ 具有 r 阶连续偏导数，$r>\max(0, m-p)$. 如果 $0\in\mathbb{R}^p$ 是 ϕ 的正则值，那末对几乎所有的 $a\in V$，0 是 $\phi_a(\cdot)=\phi(a, \cdot)$：$\mathbb{R}^m\to\mathbb{R}^p$ 的正则值.

证 设

$$M=\{(a, x)\,|\,\phi(a, x)=0,\ a\in V,\ x\in U\}.$$

根据假设 0 是 ϕ 的正则值和隐函数存在定理知，M 是 $\mathbb{R}^{q+m}$ 中的光滑流形，维数为 p.

再设 π：$\mathbb{R}^q\times\mathbb{R}^m\to\mathbb{R}^q$ 定义如下：

$$\pi(a, x)=a.$$

即 π 是投影. 如果把 π 限制在 M 上，那末它映流形 M 到 $\mathbb{R}^q$，根据沙得定理，对几乎所有的 $a\in\mathbb{R}^q$，a 是 $\pi|_M$ 的正则值，即对于 $b\in\mathbb{R}^q$，存在 $y\in\mathbb{R}^m$，使得 (b, y) 是流形 M 在点 (a, x) 处的切向量，即适合

$$D_a\phi(a, x)b+D_x\phi(a, x)y=0. \qquad (2)$$

又因为 $(a, x)\in M$，所以

$$\phi(a, x)=0.$$

由(2)得到

$$\text{Range } D_a\phi(a, x)\subset \text{Range } D_x\phi(a, x),$$

从而

$$\begin{aligned}\mathbb{R}^p &= \text{Range } D\phi(a, x)\\ &= \text{Range } D_a\phi(a, x)+\text{Range } D_x\phi(a, x)\\ &\subset \text{Range } D_x\phi(a, x)\subset \mathbb{R}^p.\end{aligned}$$

所以，对几乎所有的 $a\in V$，如果 $x\in U$ 使得

$$\phi(a, x)=0,$$

那末
$$\text{Range } D_x\phi(a, x)=\mathbb{R}^p,$$

即 0 是 $\phi_a(\cdot)$ 的正则值. ∎

现在证明

定理 4 设 $f: \mathbb{B}^n\to\mathbb{B}^n$ 具有各阶连续偏导数，那末在 $\mathbb{B}^n$ 上存在 f 的不动点，即存在 $x_0\in\mathbb{B}^n$ 满足(1).

证 设 $(a, \lambda, x)\in\mathbb{B}^n\times(-\infty, 1)\times\mathbb{B}^n$, $\mathbb{B}^n=\{x\in\mathbb{R}^n \mid \|x\|<1\}$, 置

$$\phi(a, \lambda, x)=(1-\lambda)(x-a)+\lambda(x-f(x)),$$

那末
$$\phi: \mathbb{B}^n\times(-\infty, 1)\times\mathbb{B}^n\to\mathbb{R}^n$$

具有各阶连续偏导数，并且

$$D_a\phi(\bar{a}, \bar{\lambda}, \bar{x})=-(1-\bar{\lambda})I, \quad (I\text{ 单位阵})$$

因此，如果 $\bar{\lambda}\neq 1$, $\phi(\bar{a}, \bar{\lambda}, \bar{x})=0$, 那末

$$\text{Range } D\phi(\bar{a}, \bar{\lambda}, \bar{x})\supset \text{Range } D_a\phi(\bar{a}, \bar{\lambda}, \bar{x})=\mathbb{R}^n.$$

因此，当 $\bar{\lambda}\neq 1$ 时，0 是 ϕ 的正则值.

根据定理 3, 对几乎所有的 $a\in\mathbb{B}^n$, 0 是 ϕ_a 的正则值. $\phi_a: (-\infty, 1)\times\mathbb{B}^n\to\mathbb{R}^n$, 并且

$$\phi_a(0, a)=0,$$

$$D_x\phi_a(0, a)=I,$$

根据隐函数存在定理，存在$\delta>0$，使得$x:(-\delta,\delta)\to\mathbb{R}^n$满足

$$\phi_a(\lambda, x(\lambda))=0, \quad 且 \quad x(0)=a.$$

置
$$\Gamma_a=\phi_a^{-1}(0),$$
它由一些光滑的连通分支组成，每一分支与单位圆或区间成连续一一对应。由上所述，Γ_a中含有$(0,a)$的分支不可能与单位圆一一对应。该分支的极限点(λ,x)应当位于$[0,1]\times\mathbb{B}^n$中，并且适合

$$\phi(a,\lambda,x)=0,$$

如果$0<\lambda<1$，那末

$$(1-\lambda)(x-a)+\lambda(x-f(x))=0,$$

即
$$x=(1-\lambda)a+\lambda f(x),$$

因为$a\in\mathbb{B}^n$，$f(x)\in\mathbb{B}^n$，$0<\lambda<1$，所以$x\in\mathbb{B}^n$。因此，$(\lambda,x)\in(0,1)\times\mathbb{B}^n$。这样，$\Gamma_a$中含有$(0,a)$为端点的一支在$[0,1]\times\mathbb{B}^n$上可能的极限点$(\lambda,x_0)$满足$\lambda=1$，$x_0\in\mathbb{B}^n$。现在要指出$x_0$就是$f$的不动点。事实上，由

$$\phi(a,1,x_0)=0$$

得到
$$x_0-f(x_0)=0,$$

即x_0是f的不动点。它是由Γ_a中以$(0,a)$为一端点的连通支在$(0,1]\times\mathbb{B}^n$中的极限点得到的。从而$f:\mathbb{B}^n\to\mathbb{B}^n$的不动点是存在的。 ∎

现在证明定理1.

根据魏尔斯特拉斯(Weierstrass)定理，存在一列光滑的f_k，它在$\mathbb{B}^n$上一致收敛于f，并且$f_k:\mathbb{B}^n\to\mathbb{B}^n$。对于$f_k$，存在$x_k\in\mathbb{B}^n$，满足

$$x_k-f_k(x_k)=0. \tag{3}$$

设$\{x_k\}$的收敛子列就是本身，它的极限为x_0。

因为

$$\|f_k(x_k)-f(x_0)\|\leqslant\|f_k(x_k)-f(x_k)\|+\|f(x_k)-f(x_0)\|,$$

并且当 $k\to\infty$ 时，上式右端趋于 0，所以在(3)中令 $k\to\infty$，就得到

$$x_0-f(x_0)=0.$$

即 f 的不动点 x_0 存在. ∎

附录 4　菲利浦夫引理

菲利浦夫(А. Ф. Филиппов)于一九五九年在莫斯科大学学报上讨论最优控制的存在性时，证明了下述引理.

引理　设对于 $t\in[\alpha,\ \beta]$，$Q(t)\subset\mathbb{R}^m$ 是有界闭集，并且对于 $t_0\in[\alpha,\ \beta]$，$\varepsilon>0$，存在 $\delta>0$，使得当 $|t-t_0|<\delta$，$t\in[\alpha,\ \beta]$时，$Q(t)\subseteq Q(t_0)+\varepsilon\mathbb{B}$，这里 $\mathbb{B}=\{x\,|\,\|x\|<1\}$，即 $Q(\cdot)$关于包含关系具有上半连续性. 又设 f: $[\alpha,\ \beta]\times\mathbb{R}^m\to\mathbb{R}^n$ 是连续的. 如果 $y(\cdot)$: $[\alpha,\ \beta]\to\mathbb{R}^n$ 是可积的，使得

$$y(t)\in f(t,\ Q(t))=\{x\,|\,x=f(t,\ u),\ u\in Q(t)\}\qquad(1)$$

在$[\alpha,\ \beta]$上几乎处处成立，那末存在可测的 $u(\cdot)$：$[\alpha,\ \beta]\to\mathbb{R}^m$，$u(t)\in Q(t)$，使得等式

$$f(t,\ u(t))=y(t)\qquad(2)$$

在$[\alpha,\ \beta]$上几乎处处成立.

注 1　该引理也称为可测隐函数存在定理. 事实上，(1)表明存在$u(t)\in Q(t)$，使(2)成立. 该引理说明存在 $u(\cdot)$是可测的，使(2)成立.

当 $Q(t)\equiv Q$ 不依赖 t 时，如果 Q 是有界闭集，那末 $Q(\cdot)$关于包含关系是上半连续的.

注 2　当 $Q(t)\equiv Q$ 时，如果 $f(t,\ \cdot):Q\to\mathbb{R}^n$ 连续，$f(\cdot,\ u):[\alpha,\ \beta]\to\mathbb{R}^n$ 是可积的，并且存在可积的 $\mu(\cdot):[\alpha,\ \beta]\to\mathbb{R}$，使得 $\|f(t,u)\|\leqslant\mu(t)$，$\forall u\in Q$，那末引理的结论仍成立. 即如果 $y(t)\in f(t,\ Q)$，$y(\cdot)$可积，那末存在可积的 $u(\cdot)$：$[\alpha,\ \beta]\to Q$，使得(2)在$[\alpha,\ \beta]$上几乎处处成立. 这是我们引用的主要形式.

引理的证明　因为 $f(t,\ \cdot)$：$Q(t)\to\mathbb{R}^n$ 是连续的，所以对于 $y\in f(t,\ Q(t))$，在 $f(t,\ \cdot)$映象下 y 的原象是 $Q(t)$的非空闭子集. 取出 y 的原象中第一个坐标 u^1 为最小的 $u=(u^1,\ u^2,\ \cdots,\ u^m)$；它

仍是一个闭子集，再在其中取出第二个坐标 u^2 为最小的 u，如果它仅是一个点，就中止；否则再依次取坐标 u^3，u^4，… 为最小的 u，适合

$$f(t,\ u)=y.$$

我们根据上述方法由 $y(t)\in f(t,\ Q(t))$ 来决定 $u(t)$，适合 $u(t)\in Q(t)$ 和

$$f(t,\ u(t))=y(t). \tag{3}$$

现在证明：$u(\cdot)$ 是可测的.

如果已经证得 $u(\cdot)=(u^1(\cdot),\ u^2(\cdot),\ \cdots,\ u^m(\cdot))$ 中的 $u^1(\cdot)$，$u^2(\cdot),\cdots,u^r(\cdot)$ 在 $[\alpha,\ \beta]$ 上为可测，我们要证明 $u^{r+1}(\cdot)$ 在 $[\alpha,\beta]$ 上为可测. 根据鲁津(H. H. Лузин)定理(参见夏道行等著的书[37])，对于 $\varepsilon>0$，存在闭集 $F\subset[\alpha,\ \beta]$，使得

$$\mu(F)>\beta-\alpha-\varepsilon,$$

而 $y(\cdot)$，$u^1(\cdot)$，…，$u^r(\cdot)$ 在 F 上是连续的. 现在证明：集

$$F(u^{r+1}(t)\leqslant a)=\{t\in F\,|\,u^{r+1}(t)\leqslant a\}$$

是闭集. 否则，存在 $t_k\in F$，使得

$$\lim_{k\to\infty} t_k=\bar{t},$$

$$u^{r+1}(t_k)\leqslant a,$$

$$u^{r+1}(\bar{t})>a.$$

由于 $Q(\cdot)$ 在 $[\alpha,\ \beta]$ 上关于包含关系是上半连续的，所以存在常数 $C>0$，使得当 $u\in Q(t)$，$t\in[\alpha,\ \beta]$ 时 $\|u\|\leqslant C$. 从而(3)所决定的 $u(\cdot)$ 在 $[\alpha,\ \beta]$ 上是有界的，所以 $\{u(t_k)\}$ 具有收敛的子列，我们仍记为 $\{u(t_k)\}$，设它的极限为 $\tilde{u}$. 但是

$$u(t_k)\in Q(t_k),\qquad \lim_{k\to\infty} t_k=\bar{t},$$

而 $Q(\cdot)$ 关于包含关系是上半连续的，所以

$$\tilde{u}=\lim_{k\to\infty} u(t_k)\in Q(\bar{t}).$$

又因为$u^1(\cdot)$, $u^2(\cdot)$, $\cdots$, $u^r(\cdot)$和$y(\cdot)$在F上是连续的，所以

$$u^1(\tilde{t})=\tilde{u}^1,\ \cdots,\ u^r(\tilde{t})=\tilde{u}^r,$$

$$\tilde{u}^{r+1}=\lim_{k\to\infty}u^{r+1}(t_k)\leqslant a,$$

$$u^{r+1}(\tilde{t})>a,$$

$$y(\tilde{t})=\lim_{k\to\infty}y(t_k)=\lim_{k\to\infty}f(t_k,\ u(t_k))=f(\tilde{t},\ \tilde{u}). \tag{4}$$

它表明$u^{r+1}(\tilde{t})$不是适合

$$f(\tilde{t},\ u^1(\tilde{t}),\ \cdots,\ u^r(\tilde{t}),\ u^{r+1},\ \cdots,\ u^m)=y(\tilde{t})$$

的第$r+1$个坐标u^{r+1}取最小值的，它与$u^{r+1}(\tilde{t})$的定义相矛盾。所以集合

$$F(u^{r+1}(t)\leqslant a)$$

是闭集，从而$u^{r+1}(\cdot)$在F上是可测的（参见夏道行等的书[3]）。置

$$u_F^{r+1}(t)=\begin{cases}u^{r+1}(t), & t\in F,\\ 0, & t\overline{\in}F.\end{cases}$$

那末$u_F^{r+1}(\cdot)$在$[\alpha,\ \beta]$上是可测的。当$\varepsilon\to0$时，$u_F^{r+1}(\cdot)$在$[\alpha,\ \beta]$上按测度收敛于$u^{r+1}(\cdot)$，所以$u^{r+1}(\cdot)$在$[\alpha,\ \beta]$上是可测的。

附录5 波赫纳积分

本附录假设 X 是巴拿赫空间，E 是 $\mathbb{R}$ 中的有界勒贝格可测集.

定义1 X-值函数 $x(\cdot)$：$E\to X$ 称为是弱可测的，如果对于 $f\in X^*$，$f(x(\cdot))$：$E\to\mathbb{R}$ 是勒贝格可测的；如果 $x(\cdot)$ 的值域 $\{x(s)\,|\,s\in E\}$ 是 X 中的有限集 $\{x_j\}_{j=1}^m$，且 x_j 的原象 $E_j\subset E(j=1, 2, \cdots, m)$ 是可测集，即当 $s\in E_j$ 时 $x(s)=x_j$，就称 $x(\cdot)$ 是有限

(1)推知

$$\lim_{n\to\infty} f(x_n(s)) = f(x(s)).$$

所以$f(x(\cdot))$是可测的，即$x(\cdot)$是弱可测的。

充分性．不失一般性，可以认为值域$\{x(s)|s\in E\}$本身是可分的．因此，我们可以假定X是可分的，否则用含有$x(\cdot)$的值域的线性闭子空间代替X。

首先证明$\|x(\cdot)\|$是可测的．对任一实数a，置

$$A=\{s|\|x(s)\|\leqslant a\},$$

$$A_f=\{s|f(x(s))\leqslant a\},$$

这里$f\in X^*$，$\|f\|\leqslant 1$．显然

$$A\subset A_f,$$

从而

$$A\subset \bigcap_{\|f\|\leqslant 1} A_f.$$

又，对于固定的s，如果$\|x(s)\|>0$，因为球$\{x|\|x\|\leqslant\|x(s)\|\}$是含内点的，根据附录1的定理6，存在$f_0\in X^*$，满足

$$f_0(x)\leqslant f_0(x(s)),\quad \forall\|x\|\leqslant\|x(s)\|, \tag{2}$$

$$\|f_0\|=1.$$

这时，由(2)得到

$$1=\sup\left\{\frac{f_0(x)}{\|x\|}\,\Big|\,\|x\|\leqslant\|x(s)\|\right\}=\frac{f_0(x(s))}{\|x(s)\|},$$

即

$$f_0(x(s))=\|x(s)\|.$$

所以，如果$s\in A$，则$\|x(s)\|\leqslant a$，从而$f_0(x(s))=\|x(s)\|\leqslant a$，即$s\in A_{f_0}$，因此

$$A\supset \bigcap_{\|f\|\leqslant 1} A_f.$$

从而

$$A=\bigcap_{\|f\|\leqslant 1} A_f.$$

根据附录1的定理2，我们得到$A_f=\cap A_{f_j}$，从而

$$\bigcap_{\|f\|\leqslant 1} A_f=\bigcap_{j=1}^{\infty} A_{f_j}.$$

所以 $A=\bigcap_{j=1}^{\infty}A_{f_j}$. 由于 A_{f_j} 是可测的，所以 A 可测，从而 $\|x(\cdot)\|$ 可测.

因为值域 $\{x(s)\,|\,s\in E\}$ 是可分的，所以对任何自然数 n，使该值域可以用可数个半径 $<\frac{1}{n}$ 的开球 $S_{j,n}(j=1, 2, \cdots)$ 所覆盖. 设 $S_{j,n}$ 的中心为 $x_{j,n}$，我们在前面已证 $\|x(\cdot)-x_{j,n}\|$ 是可测的，从而集 $B_{j,n}=\{s\in E\,|\,x(s)\in S_{j,n}\}$ 是可测的，且

$$E=\bigcup_{j=1}^{\infty}B_{j,n}.$$

当 $s\in B_{i,n}\setminus\bigcup_{j=1}^{i-1}B_{j,n}$ 时，置

$$x_n(s)=x_{i,n}.$$

由于 $E=\bigcup_{i=1}^{\infty}\left(B_{i,n}\setminus\bigcup_{j=1}^{i-1}B_{j,n}\right)$，所以对 $s\in E$ 有

$$\|x(s)-x_n(s)\|<\frac{1}{n}.$$

容易知道 $x_n(\cdot)$ 是强可测的，从而它的极限 $x(\cdot)$ 是强可测的. ∎

如果 $x(\cdot)$: $E\to X$ 是有限值的，且当 $s\in B_i$ 时 $x(s)=x_i(i=1, 2, \cdots, n)$，$E=\bigcup_{j=1}^{n}B_j$，我们用 $\sum_{j=1}^{n}x_j\mu(B_j)$ 来定义 $x(\cdot)$ 在 E 上的积分.

定义 3 设 $x(\cdot): E\to X$，如果存在有限值的序列 $\{x_n(\cdot)\}$，使得 $x_n(\cdot)$ 几乎处处强收敛于 $x(\cdot)$，且

$$\lim_{n\to\infty}\int_E\|x(s)-x_n(s)\|ds=0, \tag{3}$$

就称 $x(\cdot)$ 在 E 上是波赫纳可积的，并且定义 $x(\cdot)$ 的积分为

$$\int_E x(s)ds=\lim_{n\to\infty}\int_E x_n(s)ds. \tag{4}$$

首先要说明上述定义的合理性. 这是因为

$$\left\|\int_E x_n(s)ds-\int_E x_k(s)ds\right\|\leqslant\int_E\|x_n(s)-x_k(s)\|ds$$

$$\leqslant\int_E\|x_n(s)-x(s)\|ds+\int_E\|x(s)-x_k(s)\|ds \tag{5}$$

上式当 $n, k\to\infty$ 时右端趋于0，而 X 是完备的，所以(4)式右端的极限是存在的．还可以证明积分不依赖于序列 $\{x_n(\cdot)\}$ 的选择．

定理2 设 $x(\cdot)$：$E\to X$ 强可测．那末当且仅当 $\|x(\cdot)\|$ 可积时，$x(\cdot)$在 E 上是波赫纳可积的．

证 设 $\{x_n(\cdot)\}$ 是有限值的序列，几乎处处强收敛于 $x(\cdot)$．当 $x(\cdot)$波赫纳可积时，(3)式成立．由(3)及不等式

$$\int_E\big|\|x_n(s)\|-\|x_k(s)\|\big|ds$$

$$\leqslant\int_E\|x_n(s)-x(s)\|ds+\int_E\|x(s)-x_k(s)\|ds$$

推知极限 $\lim\limits_{n\to\infty}\int_E\|x_n(s)\|ds$ 存在．由不等式

$$\|x(s)\|\leqslant\|x_n(s)\|+\|x(s)-x_n(s)\|$$

及 $\|x(\cdot)\|$ 的可测性推知 $\|x(\cdot)\|$ 在 E 上是可积的．

反之，设 $\|x(\cdot)\|$ 在 E 上是可积的．置

$$y_n(s)=\begin{cases}x_n(s), & \text{当 } \|x_n(s)\|\leqslant\|x(s)\|(1+2^{-1})\text{时},\\ 0, & \text{当 } \|x_n(s)\|>\|x(s)\|(1+2^{-1})\text{时}.\end{cases}$$

那末，$y_n(\cdot)$是有限值的，满足 $\|y_n(s)\|\leqslant\|x(s)\|(1+2^{-1})$ 和 $y_n(s)$强收敛于 $x(s)$在 E 上几乎处处成立．由 $\|x(\cdot)\|$ 的可积性和

$$\|x(s)-y_n(s)\|\leqslant\|x(s)\|(2+2^{-1})$$

知 $\|x(\cdot)-y_n(\cdot)\|$ 是可积的，根据勒贝格积分的控制收敛定理，得到

$$\lim_{n\to\infty}\int_E\|x(s)-y_n(s)\|ds=0.$$

从而 $x(\cdot)$ 在 E 上是波赫纳可积的.

对波赫纳积分，利用第一章的方法，容易证明.

定理3 设 $x(\cdot): E\to X$ 是波赫纳可积的，$\lambda\in(0, 1)$，那末对于 $\varepsilon>0$，存在 $E_\lambda^\varepsilon\subset E$，$\mu(E_\lambda^\varepsilon)=\lambda\mu(E)$，使得

$$\left\|\lambda\int_E x(s)ds-\int_{E_\lambda^\varepsilon} x(s)ds\right\|<\varepsilon.$$

定理4 对于 $t\in[\alpha, \beta]$，$x(t, \cdot)$: $E\to X$ 是波赫纳可积的，并且

$$\lim_{\bar{t}\to t}\int_E\|x(\bar{t}, s)-x(t, s)\|ds=0,$$

$\lambda\in(0, 1)$，那末对于 $\varepsilon>0$，存在 $E_\lambda^\varepsilon\subset E$，满足

$$\mu(E_\lambda^\varepsilon)=\lambda\mu(E),$$

$$\left\|\lambda\int_E x(t, s)ds-\int_{E_\lambda^\varepsilon} x(t, s)ds\right\|<\varepsilon.$$

定理5 设 $x(\cdot):[\alpha, \beta]\to X$ 是波赫纳可积的，$\lambda\in(0, 1)$，那末对于 $\varepsilon>0$，存在 $E_\lambda^\varepsilon\subset[\alpha, \beta]$，满足

$$\mu(E_\lambda^\varepsilon)=(\beta-\alpha)\lambda,$$

$$\left\|\lambda\int_\alpha^t x(s)ds-\int_{[\alpha, t]\cap E_\lambda^\varepsilon} x(s)ds\right\|<\varepsilon.$$

它们的证明留给读者自行完成.

附录6　最大值原理的非光滑分析证明

史树中

从抽象的观点来看，Понтрягин 最大值原理无非是一个极值问题的一阶必要条件. 所谓极值问题的一阶必要条件，粗糙地说，是指：对于一个定义在某个带线性结构的集合上的函数，如果它在该集合的某点上达到极小值，那么函数在该点上对于任何“容许方向”上的“方向导数”都不小于零. 如果集合很正规，例如，m 维空间(对应 m 个变量的无约束最优化问题)、由多变量光滑函数的等式确定的流形(对应带等式约束的最优化问题)，而被求极值的函数又是光滑的，那么我们立即导得熟知的 Fermat 定理和 Lagrange 乘子定理. 最优控制问题的困难恰恰在于它所涉及的极值问题中，自变量变化的集合不太正规；这里被求极值的函数是控制问题的目标函数，其自变量是状态和控制，而它们的变化范围由状态方程和容许控制集等来决定. 即使目标函数对状态与控制来说都很正规，但状态方程与容许控制集合会使这个函数在一个古怪的集合上求极值；或者把控制也用状态来隐含表示时，目标函数会变成状态的古怪的函数. 这样一来，要弄清集合在一个点上的“容许方向”和函数的“方向导数”都变得不太容易. 于是在许多经典证明中都不得不另觅途径，并要用上诸如 Brouwer 不动点定理之类的工具，才能奏效. 然而，如果能适当假定最优控制问题有很正规的“容许方向”和“方向导数”，那么最大值原理的证明就会变得很简单. 例如，用动态规划方法证明最大值原理时，若能假定值函数是二次可微的，则最大值原理很自然地就可导出. 本书第四

章在值函数满足局部 Lipschitz 条件时，对最大值原理的证明也是很简单的.

非光滑分析是近二十年中发展起来的新学科．其主要目的之一就是要把那种非光滑的极值问题的一阶必要条件说清楚．这里当然不是光给出一个抽象结果就完，而是需要由这类抽象结果能导出许多能验证的有意义结果来．因此，为了适应各种问题的需要，人们不得不引入许多种"容许方向"和"方向导数"的定义，并且它们还不能互相代替(参看 Frankowska〈16〉及其所附参考文献)．非光滑分析出现后，人们自然会想到，能否以此来重新给出最大值原理的证明，但早期的一些工作并未在观念上有较大变化.

最近，Frankowska〈17〉，〈18〉在 Pallu de la Barrière [20] 的基础上，得到了最大值原理的简洁的非光滑分析证明．这些证明由于使用了非光滑分析的语言，对最大值原理作为极值问题的一阶必要条件这点揭示得更为清楚，因而也就较容易理解．尤其是在终端自由的情形，最大值原理可由无约束的非光滑极值问题的一阶必要条件立即导出；而在终端有约束情形，最大值原理同样也可由约束的非光滑极值问题的一阶必要条件导出，只是这里的有些处理就显得有点复杂．尽管如此，这些最大值原理的证明带来了某种观念上的更新，值得在这里向我国读者作介绍.

Frankowska 的证明用到的非光滑分析工具并不多．实际上只需要下列最简单情形的"容许方向"的概念.

设 K 为 $\mathbb{R}^m$ 中的任意集合．对于 $y\in\mathbb{R}^m$，定义

$$T_K(y):=\{v\in\mathbb{R}^m \mid \exists h_n\to 0^+,\ v_n\to v,\ y+h_nv_n\in K\}. \tag{1}$$

它称为 K 在点 y 的相依锥(contingent cone)．这是 K 在点 y 处的某种意义下的"容许方向"全体．这个概念容易推广到 Banach 空间等一般情形．有了这个概念以后，对于定义在 K 的邻域上的

光滑函数 F，如果它在 $y\in K$ 上达到极小值，那么其一阶必要条件显然为

$$\forall v\in T_k(y),\quad \langle\nabla F(y),\, v\rangle\geqslant 0. \tag{2}$$

现在我们来考虑下列终端自由的最优控制问题.

(P1) $$\min g(x(T)),$$

其中 $g:\mathbb{R}^m\to\mathbb{R}$ 为连续可微函数，而 $x(\cdot):[0,\ T]\to\mathbb{R}^m$ 为满足下列终端自由的状态方程的绝对连续映射:

$$\begin{cases}\dot{x}(t)=f(x(t),\ u(t)) & \text{a.e.}\\ x(0)=x_0 \\ u(t)\in U\end{cases} \tag{3}$$

这里 U 是控制区域，它可以是有限维空间中的任意集合，甚至可一般地假定为某个距离空间; $f:\mathbb{R}^m\times U\to\mathbb{R}^m$ 连续; $u:[0,\ T]\to U$ 可测. 下面的证明对 $u(\cdot)$ 按段连续也成立.

如果对于任何 $t\in[0,\ T]$，令

$$R(t):=\{x(t)\,|\,\text{对于某个}\ u(\cdot),\ x(\cdot)\text{是(3)的解.}\} \tag{4}$$

它称为系统(3)的轨线在时刻 t 的能达集，那么由条件(1)，(2)，我们立即可得最优控制问题(P1)的一阶必要条件为

$$\forall v\in T_{R(T)}(x(T)),\quad \langle\nabla g(x(T)),\ v\rangle\geqslant 0. \tag{5}$$

引入 $\mathbb{R}^m$ 的锥 C 的对偶锥 C^- 如下.

$$C^-:=\{p\in\mathbb{R}^m\,|\,\forall q\in C,\ \langle p,\ q\rangle\leqslant 0\}, \tag{6}$$

则条件(5)又可写成

$$-\nabla g(x(T))\in T_{R(T)}(x(T))^-. \tag{7}$$

问题归结为由(7)究竟可导出什么结论. 它与最大值原理能挂上钩的关键在于下列两条定理.

定理 1 (Frankowska〈17〉) 设 $t_0\in[0,T]$，$\bar{u}(\cdot):[0,\ T]\to U$ 可测，$\bar{x}(\cdot):[0,\ T]\to\mathbb{R}^m$ 是系统(3)对应 $\bar{u}(\cdot)$ 的绝对连续轨线. 假定 $\bar{u}(\cdot)$，$\bar{x}(\cdot)$ 满足

H1) 对于几乎所有的 $t\in[0,T]$，f 在 $(\bar{x}(t),\bar{u}(t))$ 处对 x 可微。

H2) 存在 $\varepsilon>0$，$k(\cdot)\in L^1(0, T; \mathbb{R})$，使得对于任何 $u\in U$，$t\in[0, T]$，f 满足

$$\forall x_1, x_2\in B(\bar{x}(t), \varepsilon), \ \|f(x_1, u)-f(x_2, u)\|\leqslant k(t)\|x_1-x_2\|,$$

$$B(\bar{x}(t), \varepsilon):=\{y\in\mathbb{R}^m \mid \|y-\bar{x}(t)\|<\varepsilon\}.$$

那么，对于任何 $v\in T_{R(t_0)}(\bar{x}(t_0))$，下列变分方程

$$\begin{cases}\dfrac{d}{dt}\delta x(t)=\dfrac{\partial}{\partial x}f(\bar{x}(t), \bar{u}(t))\delta x(t), \\ \delta x(t_0)=v,\end{cases} \tag{8}$$

在 $[t_0, T]$ 中的解 $\delta x(\cdot)$ 满足

$$\forall t\in[t_0, T], \quad \delta x(t)\in T_{R(t)}(\bar{x}(t)). \tag{9}$$ ▌

回顾相依锥和能达集的定义(1)和(4)，我们可以看出定理 1 不过是第四章 § 4 的结果的另一种说法。但是，有了这两个概念以后，它被表达为能达集在系统状态处的相依锥对变分方程的不变性。

推论 设 $\psi(\cdot)$ 是共轭方程

$$-\dot{\psi}(t)=\left(\frac{\partial}{\partial x}f(\bar{x}(t), \bar{u}(t))\right)^T\psi(t) \tag{10}$$

的解，且满足

$$\psi(t_0)\in T_{R(t_0)}(\bar{x}(t_0))^-, \tag{11}$$

则

$$\forall t\in[0, t_0], \quad \psi(t)\in T_{R(t)}(\bar{x}(t))^-. \tag{12}$$

特别是

$$\psi(T)\in T_{R(T)}(\bar{x}(T))^-\Rightarrow\forall t\in[0, T], \ \psi(t)\in T_{R(T)}(\bar{x}(T))^-. \tag{13}$$

证明 对(8)和(10)应用第一章 § 5 公式(13)可得

$$\langle\psi(t), \delta x(t)\rangle=\langle\psi(t_0), \delta x(t_0)\rangle, \tag{14}$$

对任何 $v\in T_{R(t)}(\bar{x}(t))$，$t\leqslant t_0\leqslant T$，在(8)，(9)中对调 t 与 t_0 的位

置，则由(9)有 $\delta x(t_0)\in T_{R(t_0)}(\bar{x}(t_0))$，从而由(11)，(14)可得

$$\langle\psi(t),\ v\rangle=\langle\psi(t_0),\ \delta x(t_0)\rangle\leqslant 0$$

因为 v 是任意的，所以 $\psi(t)\in T_{R(t)}(\bar{x}(t))^-$. ∎

定理 2 (Pallu de la Barrière [20]) 在定理 1 的假定下，设 $t\in(0,\ T]$ 为使得 $\bar{x}(\cdot)$ 的左导数 $\bar{x}'_-(t)=f(\bar{x}(t),\ \bar{u}(t))$ 的点，则

$$\forall u\in U,\ f(\bar{x}(t),\ u)-f(\bar{x}(t),\ \bar{u}(t))\in T_{R(t)}(\bar{x}(t)). \tag{15}$$

证明 由假定，我们有

$$\bar{x}(t)=\bar{x}(t-h)+hf(\bar{x}(t),\ \bar{u}(t))+o(h), \tag{16}$$

其中 $h>0$ 充分小. 固定 $u\in U$，并令 $x_h(\cdot)$ 为下列方程的解

$$\begin{cases}\dot{x}_h(t)=f(x_h(t),\ u),\\ x_h(t-h)=\bar{x}(t-h),\end{cases} \tag{17}$$

则
$$x_h(t')=\bar{x}(t-h)+\int_{t-h}^{t'}f(x_h(s),\ u)ds$$

因此，对于充分小的 $h>0$，由 H2)，我们有

$$\begin{aligned}&\|x_h(t)-\bar{x}(t-h)-hf(\bar{x}(t-h),\ u)\|\\ &=\left\|\int_{t-h}^{t}\{f(x_h(s),\ u)-f(\bar{x}(t-h),\ u)\}ds\right\|\\ &\leqslant\int_{t-h}^{t}k(s)\|x_h(s)-\bar{x}(t-h)\|ds\\ &\leqslant\int_{t-h}^{t}k(s)\int_{t-h}^{s}\|f(x_h(s'),\ u)\|ds'ds=o(h).\end{aligned} \tag{18}$$

由(16)，(18)和 H1)，立即得到

$$\begin{aligned}x_h(t)&=\bar{x}(t-h)+hf(\bar{x}(t-h),\ u)+o(h)\\ &=\bar{x}(t-h)+hf(\bar{x}(t),\ u)+o(h)\\ &=\bar{x}(t)+h[f(\bar{x}(t),\ u)-f(\bar{x}(t),\ \bar{u}(t))]+o(h).\end{aligned}$$

因为由(17)易知 $x_h(t)\in R(t)$，由相依锥的定义(1)，即得(15)成立. ∎

推论(最大值原理) 设 $\psi(\cdot)$：$[0,\ T]\to\mathbb{R}^n$ 为共轭方程(10)

的解，且满足

$$\psi(T)\in T_{R(T)}(\bar{x}(T))^{-},\tag{19}$$

那么对于几乎所有的 $t\in[0, T]$，

$$\langle\psi(t), f(\bar{x}(t), \bar{u}(t))\rangle=\max_{u\in U}\langle\psi(t), f(\bar{x}(t), u)\rangle\tag{20}$$

证明由(13)，(15)立即可得. ∎

至此，我们已基本完成对终端自由的最大值原理的非光滑分析证明. 实际上，这里并未对控制的最优性作任何假定，而仅指出对于任意的控制 $\bar{u}(\cdot)$ 及其对应的轨线 $\bar{x}(\cdot)$，控制区域 U 和对应的共轭状态轨线 $\psi(\cdot)$ 都与能达集在轨线 $\bar{x}(\cdot)$ 能达点处的相依锥有某种不变关系. 至于通常所说的最大值原理，只是把这一关系再与最优性的必要条件联系起来而已. 例如对于问题(P1)，由(7)和(19)，令

$$\psi(T)=-\nabla g(\bar{x}(T))\tag{21}$$

就得到通常所说的最大值原理的所有关系，即(10)，(20)，(21).

对于更一般的终端自由的最优控制问题也可类似讨论. 例如，对于下列问题

$$\text{(P2)}\qquad\begin{cases}\min\ \phi(x(0), x(T))+\int_0^T L(x(t), u(t))dt,\\ \dot{x}(t)=f(x(t), u(t)),\\ x(0)\in K,\ u(t)\in U.\end{cases}$$

其最大值原理为

$$\begin{cases}-\dot{\psi}(t)=\left(\frac{\partial}{\partial x}f(\bar{x}(t),\bar{u}(t))\right)^T\psi(t)-\frac{\partial}{\partial x}L(\bar{x}(t),\bar{u}(t)),\ \text{a.e.}\\ \psi(0)\in\frac{\partial}{\partial x_1}\phi(\bar{x}(0), \bar{x}(T))+T_K(\bar{x}(0))^{-},\\ \psi(T)=-\frac{\partial}{\partial x_2}\phi(\bar{x}(0), \bar{x}(T)),\\ \langle\psi(t), f(\bar{x}(t), \bar{u}(t))\rangle-L(\bar{x}(t), \bar{u}(t))\\ \quad=\max_{u\in U}\{\langle\psi(t), f(\bar{x}(t), u)\rangle-L(\bar{x}(t), u)\}.\quad\text{a.e.}\end{cases}$$

这里 ϕ, L 都要求有一定的正则性. 与经典的结果不同之处在于现在对于初值的约束集 K 可很自然地引入 $T_K(\bar{x}(0))^-$ 项.

对 T 不固定的自由终端问题仍可用同样方式来讨论. 然而, 对于终端固定, 或更一般的终端有约束的问题, 上述讨论就失效了. 问题在于我们不能再利用类似(5)—(7)那样的关系, 而应该代之以某种 Lagrange 乘子型的条件. 下面为明确起见, 我们考虑下列有终端约束的最优控制问题.

$$
\text{(P3)} \qquad \begin{cases} \min \phi(x(0),\ x(T)), \\ \dot{x}(t)=f(x(t),\ u(t)), \quad u(t)\in U, \\ x(0)\in K_0, \quad x(T)\in K_T, \end{cases}
$$

其中 $\phi: \mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}$ 连续可微, K_0, $K_T\subset\mathbb{R}^m$ 为两个闭集.

记所有的容许控制全体为

$$
\mathscr{U}:=\{u(\cdot):[0,\ T]\to U \mid u \text{ 可测}\}. \tag{22}
$$

在 $\mathscr{U}$ 上定义

$$
d(u(\cdot),\ v(\cdot)):=\mu(\{t\in[0,\ T] \mid u(t)\neq v(t)\}), \tag{23}
$$

这里 μ 表示 $[0,\ T]$ 上的 Lebesgue 测度. 则可以证明 d 是 $\mathscr{U}$ 上的距离, 且 $\mathscr{U}$ 关于 d 是完备距离空间. 再定义 $\Phi: \mathbb{R}^m\times\mathscr{U}\to\mathbb{R}^m\times\mathbb{R}^m$ 为

$$
\Phi(x_0,\ u(\cdot))=(x(0),\ x(T)),
$$

其中 $x(\cdot)$ 是(3)的解. 令 $\mathscr{V}:=\mathbb{R}^m\times\mathscr{U}$, $X:=\mathbb{R}^m\times\mathbb{R}^m$, $K=K_0\times K_T\subset X$, $J:=\phi\circ\Phi:\mathscr{V}\to\mathbb{R}$. 则问题(P3)可表达为下列抽象的约束极值问题:

$$
(\mathscr{P}) \qquad \begin{cases} \min J(v), \\ v\in\mathscr{V}, \\ \Phi(v)\in K\subset X. \end{cases}
$$

这里 J 是完备距离空间 $\mathscr{V}$ 上的连续函数, Φ 是 $\mathscr{V}$ 到(有限维) Hilbert 空间 X 的连续映射, K 是 X 中的闭集. 我们首先要对问题($\mathscr{P}$)提出一条非常一般的 Lagrange 乘子型的定理. 这里的

困难不但在于 J, Φ, K 等不一定是正则的，还在于 $\mathscr{V}$ 是个距离空间，它不一定有与距离协调的线性结构，正如对 $\mathscr{U}$ 所提出的距离那样．为此，需要提出距离空间到 Hilbert 空间的映射的变分的概念．

设 F: $\mathscr{V}\to X$ 为任意映射．定义 F 在点 $v\in\mathscr{V}$ 上的一阶相依变分为

$$V_F(v):=\{y\in X\,|\,\exists h_n\to0^+,\ y_n\to y,\ F(v)+h_ny_n\in F(B_{h_n}(v))\},$$

其中 $B_h(v)=\{u\in\mathscr{V}\,|\,d(u,\ v)<h\}$．当 $\mathscr{V}$ 为 Banach 空间，F 连续可微时，则由 $B_h(v)=v+hB$（B 为 $\mathscr{V}$ 中的以原点为中心的单位开球），$F(v+hB)-F(v)=hF'(u)B+o(h)$，立即可得 $V_F(v)=\overline{F'(v)B}$．问题在于我们这里的 $\mathscr{V}$ 不是 Banach 空间，故需要有这样的推广．

定理 3 (Frankowska〈18〉) 设 $\bar{v}\in\mathscr{V}$ 是问题($\mathscr{P}$)的解．那末存在序列 $\{v_k\}\subset\mathscr{V}$, $v_k\to\bar{v}$, $\{x_k\}\subset K$, $x_k\to\Phi(\bar{v})$ 以及 $\lambda_k\geqslant0$, $\zeta_k\in X$, $\lambda_k^2+\|\zeta_k\|^2=1$, 对于任何 $k=1,\ 2,\ \cdots$, 满足

$$\forall y\in T_k(x_k),\qquad\qquad \langle\zeta_k,\ y\rangle\leqslant\|y\|/2k, \tag{24}$$

$$\forall(w,\ z)\in V_{(J,\Phi)}(v_k),\qquad \lambda_kw+\langle\zeta_k,\ z\rangle\geqslant-1/2k. \tag{25}$$

证明 对于任何 $k\geqslant1$, 定义

$$\begin{cases} f_k\colon \mathscr{V}\to\mathbb{R},\ f_k(v)=\max\left\{0,\ J(v)-J(\bar{v})+\dfrac{1}{k^2}\right\}, \\ F_k\colon \mathscr{V}\times K\to\mathbb{R},\ F_k(v,\ x)=f_k^2(v)+\|\Phi(v)-x\|^2. \end{cases} \tag{26}$$

那么 F_k 是完备距离空间 $\mathscr{V}\times K$ 上的非负连续函数，且

$$F_k(\bar{v},\ \Phi(\bar{v}))=1/k^4.$$

对 F_k 和点$(\bar{v},\ \Phi(\bar{v}))$应用 Ekeland 变分原理，可求得 $v_k\in\mathscr{V}$, $x_k\in K$, 满足

$$\begin{aligned} &d(v_k,\ \bar{v})+\|\Phi(\bar{v})-x_k\|\leqslant1/k, \\ &F_k(v_k,\ x_k)\leqslant F_k(\bar{v},\ \Phi(\bar{v}))=1/k^4, \end{aligned} \tag{27}$$

且

$$\forall (v,\ x)\in \mathscr{V}\times K,$$
$$F_k(v,\ x)\geqslant F_k(v_k,\ x_k)-\frac{1}{k^3}(d(v,\ v_k)+\|x-x_k\|). \tag{28}$$

此外，由 $\bar{v}$ 是问题($\mathscr{P}$)的解，还可得到

$$F_k(v_k,\ x_k)>0. \tag{29}$$

在(28)中令 $v=v_k,\ x=x_k+hy'\in K$，则得

$$\|\Phi(v_k)-x_k-hy'\|^2-\|\Phi(v_k)-x_k\|^2\geqslant -\frac{h}{k^3}\|y'\|.$$

如果 $y\in T_K(x_k)$，则存在 $h_n\to 0^+,\ y_n\to y$，使得 $x_k+h_ny_n\in K$，从而

$$\lim_{n\to\infty}\frac{1}{h_n}\{\|\Phi(v_k)-x_k-h_ny_n\|^2-\|\Phi(v_k)-x_k\|^2\}$$
$$=2\langle\Phi(v_k)-x_k,\ -y\rangle\geqslant -\|y\|/k^3$$

即

$$\forall y\in T_K(x_k),\quad \langle\Phi(v_k)-x_k,\ y\rangle\leqslant\|y\|/2k^3. \tag{30}$$

另一方面，在(28)中，令 $x=x_k$，则得

$$[f_k^2(v)+\|\Phi(v)-x_k\|^2]-[f_k^2(v_k)+\|\Phi(v_k)-x_k\|^2]$$
$$\geqslant -\frac{1}{k^3}d(v,\ v_k). \tag{31}$$

如果 $(w,\ z)\in V_{(J,\Phi)}(v_k)$，则存在 $h_n\to 0^+,\ (w_n,\ z_n)\to(w,\ z)$，使得 $(J(v_k),\ \Phi(v_k))+h_n(w_n,\ z_n)\in(J,\ \Phi)(B_{h_n}(v_k))$，即存在 $v_{kn}\in B_{h_n}(v_k)$ 满足 $J(v_k)+h_nw_n=J(v_{kn}),\ \Phi(v_k)+h_nz_n=\Phi(v_{kn})$

这样，在(31)中取 $v=v_{kn}$，可得

$$\max\left\{0,\ J(v_k)+h_nw_n-J(\bar{v})+\frac{1}{k^2}\right\}^2$$
$$-\max\left\{0,\ J(v_k)-J(\bar{v})+\frac{1}{k^2}\right\}^2$$
$$+\|\Phi(v_k)+h_nz_n-x_k\|^2-\|\Phi(v_k)-x_k\|^2$$
$$\geqslant -d(v_{kn},\ v_k)/k^3\geqslant -\frac{h_n}{k^3},$$

两端除以 $2h_n$，并令 $n\to\infty$，即得

$$\forall (w, z)\in V_{(J,\Phi)}(v_k),\quad f_k(v_k)w+\langle \Phi(v_k)-x_k, z\rangle\geqslant -1/2k^3. \tag{32}$$

最后，由(26)，(27)，(29)可知

$$1/k^4\geqslant F_k(v_k, x_k)=f_k^2(v_k)+\|\Phi(v_k)-x_k\|^2>0.$$

令

$$\lambda_k=f_k(v_k)/\sqrt{F_k(v_k, x_k)},\quad \zeta_k=(\Phi(v_k)-x_k)/\sqrt{F_k(v_k, x_k)},$$

则由(27)，(30)，(32)可知定理成立. ∎

定理 3 是个抽象结果. 对于我们的具体问题(P3)，$J=\phi\circ\Phi$. 这时，(25)式可代替为

$$\forall z\in V_\Phi(v_k),\quad \langle \lambda_k\nabla\phi(v_k)+\zeta_k, z\rangle\geqslant -1/2k. \tag{33}$$

因为容易验证，$(\nabla\phi(v_k)z, z)\in V_{(J,\Phi)}(v_k)$.

为了更好地表达定理 3，我们引进集合列的 Kuratowski 下极限的概念. 设 $\{A_k\}$ 为距离空间 E 中的一个集合列. 则

$$\liminf_{k\to\infty} A_k=A:=\{x\in E\mid \exists x_k\in A_k,\ \lim_{k\to\infty} x_k=x\}$$

称为 $\{A_k\}$ 的 Kuratowski 下极限. 对集合族的下极限也可类似定义. 在非光滑分析中所应用的各种"容许方向"集(切向锥)中，有以下一个重要结果.

引理(Cornet，参看[12]409 页，定理 7)　设 K 为有限维空间 X 中的闭集. 那么对于任何 $\bar{x}\in K$，有

$$\liminf_{x\to_K \bar{x}} T_K(x)=C_K(\bar{x}):=\{v\in X\mid \forall h_n\to 0^+,\ \forall x_n\xrightarrow{K}\bar{x},\ \exists v_n\to v,\ x_n+h_nv_n\in K\}$$

这里 $x\xrightarrow{K}\bar{x}$ 表示 x 在 K 中趋向于 $\bar{x}$. $C_K(\bar{x})$ 称为 K 在 $\bar{x}$ 处的 Clarke 切向锥. ∎

有了这些概念以后，定理 3 可改述为.

定理 3′　设 $\bar{v}\in\mathscr{V}$ 是问题($\mathscr{P}$)的解. 那么存在 $\lambda\geqslant 0$, $\zeta\in X$,

$\lambda^2+\|\zeta\|^2=1$，满足

$$\forall y\in C_K(\Phi(\bar v)),\quad \langle\zeta,\ y\rangle\leqslant 0,$$

即

$$\zeta\in N_K(\Phi(\bar v)):=C_K(\Phi(\bar v))^-. \tag{34}$$

$$\forall(w,\ z)\in\liminf_{v\to\bar v}V_{(J,\Phi)}(v),\quad \lambda w+\langle\zeta,\ z\rangle\geqslant 0. \tag{35}$$

如果 $J=\phi\circ\Phi$，ϕ 连续可微，则(35)可代替为

$$\forall z\in\liminf_{v\to\bar v}V_\Phi(v),\quad \langle\lambda\nabla\phi(\bar v)+\zeta,\ z\rangle\geqslant 0. \tag{36}$$ ▮

证明只需注意到我们可对以前求得的点列 $\{(\lambda_k,\ \zeta_k)\}$ 选取一收敛子列，而该子列的极限 $(\lambda,\ \zeta)$ 即满足(34)—(36)．

下面的问题在于首先要指出，当初值 $\bar x_0\in K_0$ 和控制 $\bar u(\cdot)\in\mathscr U$ 构成的 $\bar x=(\bar x_0,\ \bar u(\cdot))\in\mathscr V=\mathbb R^m\times\mathscr U$ 是问题(P3)及其一般化$(\mathscr P)$的解时，$\liminf_{v\to\bar v}V_\Phi(v)$是什么．事实上，我们有下列一般定理．

定理 4(Frankowska 〈18〉)　设 $(\bar x(\cdot),\ \bar u(\cdot))$ 为系统(3)的解，且 H1)，H2)满足．那么对于任何 $y\in\mathbb R^m$，$\|y\|\leqslant 1$，变分方程

$$\begin{cases}\dfrac{d}{dt}\delta x(t)=\dfrac{\partial}{\partial x}f(\bar x(t),\ \bar u(t))\delta x(t),\\ \delta x(0)=y,\end{cases} \tag{37}$$

的解 $\delta x(\cdot)$满足

$$(y,\ \delta x(T))\in V_\Phi(\bar x(0),\ \bar u(\cdot)), \tag{38}$$

而对于任何 $t_0\in(0,\ T]$，其上 $\dot{\bar x}'_-(t_0)=f(\bar x(t_0),\ \bar u(t_0))$，以及$u\in U$，下列变分方程

$$\begin{cases}\dfrac{d}{dt}\delta x(t)=\dfrac{\partial}{\partial x}f(\bar x(t),\ \bar u(t))\delta x(t)\\ \delta x(t_0)=f(\bar x(t_0),\ u)-f(\bar x(t_0),\ \bar u(t_0))\end{cases} \tag{39}$$

的解 $\widetilde{\delta x}(\cdot)$ 也满足

$$(0,\ \widetilde{\delta x}(T))\in V_\Phi(\bar x(0),\ \bar u(\cdot)). \tag{40}$$

如果再假设

H3) $\forall u\in U$, $\frac{\partial}{\partial x}f(\cdot, u)$在$\bar{x}([0, T])$的邻域中连续. 那末(38), (40)的右端都可代替为 $\liminf_{x\to\bar{x}(0), u(\cdot)\to\bar{u}(\cdot)} V_x(x, u(\cdot))$. ∎

这一定理完全是定理1, 定理2的推广, 其证明也几乎一样. 后半段则只是利用了一些连续性而已.

有了定理3′和定理4以后, 我们就很容易导出关于问题(P3)的最大值原理如下.

定理5(Frankowska〈18〉) 设$(\bar{x}(\cdot), \bar{u}(\cdot))$为问题(P3)的解, 且H1), H2), H3)满足, 那末存在$\lambda\geqslant 0$和共轭方程(10)的解$\psi(\cdot)$, 且$\lambda^2+\|\psi(\cdot)\|^2\neq 0$, 满足

$$(\psi(0), -\psi(T))\in\lambda\nabla\phi(\bar{x}(0), \bar{x}(T)) + N_{K_0}(\bar{x}(0))\times N_{K_T}(\bar{x}(T)), \tag{41}$$

以及对于几乎所有的$t\in[0, T]$

$$\langle\psi(t), f(\bar{x}(t), \bar{u}(t))\rangle=\max_{u\in U}\langle\psi(t), f(\bar{x}(t), u)\rangle, \tag{20}$$

其中

$$N_{K_0}(\bar{x}0))\times N_{K_T}(\bar{x}(T))=N_{K_0\times K_T}(\bar{x}(0), \bar{x}(T))$$
$$=C_{K_0\times K_T}(\bar{x}(0), \bar{x}(T))^-=C_{K_0}(\bar{x}(0))^-\times C_{K_T}(\bar{x}(T))^-$$

证明 事实上, 由定理3′, 存在$\lambda\geqslant 0$和

$$\zeta=(\zeta_0, \zeta_T)\in N_{K_0}(\bar{x}(0))\times N_{K_T}(\bar{x}(T)),\ \lambda^2+\|(\zeta_0, \zeta_T)\|^2=1,$$

使得

$$\forall z\in\liminf_{x\to\bar{x}(0), u(\cdot)\to\bar{u}(\cdot)} V_x(\bar{x}(0), \bar{u}(\cdot)),$$
$$\langle\lambda\nabla\phi(\bar{x}(0), \bar{x}(T))+(\zeta_0, \zeta_T), z\rangle\geqslant 0 \tag{42}$$

由定理4, 由此可得, 对于任何$y\in\mathbb{R}^n$, $\|y\|\leqslant 1$, (37)的解$\delta x(\cdot)$满足

$$\langle\lambda\nabla\phi(\bar{x}(0), \bar{x}(T))+(\zeta_0, \zeta_T), (y, \delta x(T))\rangle\geqslant 0. \tag{43}$$

设$\psi(\cdot)$为共轭方程(10)的满足下列条件的解:

$$-\psi(T)=\lambda\frac{\partial}{\partial x_2}\phi(\bar{x}(0),\bar{x}(T))$$

$$+\zeta_T\in\lambda\frac{\partial}{\partial x_2}\phi(\bar{x}(0),\ \bar{x}(T))+N_{K_T}(\bar{x}(T)). \tag{44}$$

由(42)可得

$$\left\langle\lambda\frac{\partial}{\partial x_1}\phi(\bar{x}(0),\ \bar{x}(T))+\zeta_0,\ \delta x(0)\right\rangle+\langle-\psi(T),\ \partial x(T)\rangle\geqslant 0,$$

但 $\langle\psi(T),\ \delta x(T)\rangle=\langle\psi(0),\ \delta x(0)\rangle$，从而

$$\left\langle\lambda\frac{\partial}{\partial x_1}\phi(\bar{x}(0),\bar{x}(T))+\zeta_0-\psi(0),\ \delta x(0)\right\rangle$$

$$=\left\langle\lambda\frac{\partial}{\partial x_1}\phi(\bar{x}(0),\ \bar{x}(T))+\zeta_0-\psi(0),\ y\right\rangle\geqslant 0.$$

由 y 可取单位球中任何值，即得

$$\psi(0)=\lambda\frac{\partial}{\partial x_1}\phi(\bar{x}(0),\ \bar{x}(T))$$

$$+\zeta_0\in\lambda\frac{\partial}{\partial x_1}\phi(\bar{x}(0),\ \bar{x}(T))+N_{K_0}(\bar{x}(0)). \tag{45}$$

此外，再由定理 4，(42)，(44)和(45)，可得对于几乎所有的 $t\in[0,\ T]$和任何 $u\in U$，有

$$\langle-\psi(T),\ \delta\bar{x}(T)\rangle=\langle-\psi(t),\ f(\bar{x}(t),u)-f(\bar{x}(t),\bar{u}(t))\rangle\geqslant 0$$

即(20)成立. ∎

这样，我们就完成了终端有约束情形的最大值原理的证明．这个证明在思路上与终端自由情形是一致的，只是在有约束时，相依锥的讨论被代替为相依变分的讨论．在数学工具方面，没有应用Brouwer不动点定理、凸集分离定理等．但为证明一般的Lagrange乘子型定理 3 和定理 3′，需要利用非光滑分析的 Ekeland 变分原理．这里还出现了新的非光滑分析概念：Clarke 切向锥及其对偶锥，后者称为法向锥($N_{K_0}(\bar{x}(0))$，$N_{K_T}(\bar{x}(T))$)．但这是由于我们考虑的问题(P3)是比终端固定更一般的终端有约束问题．如果

$K_0=\{x_0\}$, $K_T=\{x_T\}$, 则

$$N_{K_0}(\bar{x}(0))=N_{K_T}(\bar{x}(T))=\{0\},$$

定理5就回到熟知的结果，并且这些概念也都可以不在证明中出现。

参 考 书 籍

[1] 钱学森，工程控制论，科学出版社，北京，一九五八年.

[2] 钱学森、宋健，工程控制论(修订版)，科学出版社，北京，一九八〇年.

[3] 夏道行、吴卓人、严绍宗，实变数函数论与泛函分析概要，上海科学技术出版社，上海，一九六三年.

[4] 关肇直，泛函分析讲义，高等教育出版社，北京，一九五九年.

[5] 关肇直、张恭庆、冯德兴，线性泛函分析入门，上海科学技术出版社，上海，一九七九年.

[6] 金福临、李训经等，常微分方程，上海科学技术出版社，上海，一九八四年.

[7] 李训经、孙莱祥、陈有根，计算机应用中的控制理论，复旦大学出版社，上海，一九八八年.

[8] 宫锡芳，最优控制问题的计算方法，科学出版社，北京，一九七九年.

[9] 邓乃扬等，无约束最优化计算方法，科学出版社，北京，一九八二年.

[10] 庞特里雅金等著，陈祖浩等译，最佳过程的数学理论，上海科学技术出版社，上海，一九六五年.

[11] 伯科维茨著，贺建勋等译，最优控制理论，上海科学技术出版社，上海，一九八五年.

[12] **Aubin, J. P. and, Ekeland, I.** Applied Nonlinear Analysis, **Wiley-Interscience**, 1984.

[13] **Bellman, R.**, Dynamic Programming, **Princeton Univ. Press**, Princeton, New Jersey, 1957.

[14] **Bellman, R., Glicksberg, I. and Gross, O.**, Some Aspects of the Mathematical Theory of Control Processes, **The RAND Corp.**, R 313, 1958.

[15] **Clarke, F. H.**, Optimization and Nonsmooth Analysis, **John Wiley & Sons**, New York, 1983.

[16] **Hille E. and Phillips R. S.**, Functional **Analysis** and Semi-groups, **American Mathematical Society Colloquium Publications.** Vol. XXXI, 1957.

[17] **Lee, E. B. and Markus, L.**, Foundations of Optimal Control Theory, **Wiley**, New York, 1967.

[18] **Lions, J. L.**, Optimal Control of Systems governed by Partial Differential Equations, ,**Springer-Verlag**, New York, 1971.

[19] **Polak E.**, Computation Methods in Optimization: A Unified Approach, **Academic Press**, New York and London, 1971.

[20] **Pallu R. de la Barrière**, Optimal Control Theory, **W. B. Saunders Company**, Philadaphia, 1967.

[21] **Yosida K.**, Functional Analysis, **Springer-Verlag**, 1978.

参 考 文 献

〔1〕 宋健、韩京清，线性最速系统的分析与综合，数学进展，一九六二年.

〔2〕 李训经、谢惠民等，自动调节系统的稳定性和快速最佳控制，数学论文集，复旦大学数学研究所，一九六四年.

〔3〕 李训经、姚允龙，分布参数系统的时间最优控制，中国科学，一九八〇年第七期.

〔4〕 雍炯敏，对 Lindenstrauss "A short proof of Liapounoff's convexity theorem"一文的注记，数学研究与评论，一九八四年.

〔5〕 陈祖浩，最优过程罚函数方法的数学理论，数学年刊，3(3)，一九八二年.

〔6〕 陈祖浩，解受限最优控制问题的混合罚函数方法，控制理论与应用，创刊号，一九八四年.

〔7〕 陈祖浩，求解非线性控制系统最优控制的罚函数方法，自动化学报，11(4)，一九八五年.

〔8〕 陈祖浩，最优控制罚函数方法的统一理论，高校应用数学学报，一卷2期，一九八六年.

〔9〕 陈祖浩，约束最优化问题的罚函数方法，数学年刊，7A(4)，一九八六年.

〔10〕 **Balakrishnan, A. V.**, On a New Computing Technique in Optimal Control, **SIAM J. Control**, 6(2), 1968.

〔11〕 **Bushaw, D. W.**, Optimal Discontinuous Forcing Terms, in **Contributions to the Theory of Nonlinear Oscillations**,, Vol. 4, Princeton Univ. Press, Princeton, New Jersey, 1958.

〔12〕 **Chen, Zuhao**, The Optimal Control of Bounded Phase Coordinate and Penalty Functions Methods, in **Proc. of Biteleral Meeting on Control Systems**, Science Press, Beijing, China, Gordon & Breach, New York, 1982.

〔13〕 **Cullum, J.**, Penalty Functions and Nonconvex Continuous Optimal Control Problems, in **Computing Methods in Optimization.**

Vol. 2 (Zadeh, L. A., Neustadt, L. W. and Balakrishnan, A. V. eds), Academic Press, New York, 1969.

〈14〉 **Ekeland, I.,** On the Variational Principle, **J. Math. Anal. Appl.,** Vol. 47, 1974.

〈15〉 **Ekeland, I.,** Nonconvex Minimization Problems, **Bull. Amer. Math. Soc.** (New Series) 1, 1979.

〈16〉 **H. Frankowska,** The maximum principle for an optimal solution to a differential inclusion with end points constraints, **SIAM, J. Control and Optimization**, 25(1987), 145—157.

〈17〉 **H. Frankowska,** A geometric proof of Pontryagin's maximum principle, **preprint**, 1987.

〈18〉 **H. Frankowska,** Nonsmooth Analysis and Control, **preprint,** 1987.

〈19〉 **Kalman, R. E. and Bertram, J. E.,** Control System Analysis and Design via the Second Method of Lyapunov, **Trans. ASME, Ser. D.,** Vol. 82, 1960.

〈20〉 **Kalman, R. E.,** Contributions to the Theory of Optimal Control, **Bol. Soc. Mat. Mex.,** Vol. 5, 1960.

〈21〉 **Красовский, Н. Н.,** К теории оптимального регулирования, **Автоматика и Телемеханика**, т. 18, 1957.

〈22〉 **LaSalle, J. P.,** The Time Optimal Control Problem, in **Contributions to the Theory of Nonlinear Oscillations**, Vol. 5, Princeton Univ. Press, Princeton, New Jersey. 1959.

〈23〉 **Li, Xunjing,** Maximum Principle of Optimal Periodic Control for Functional Differential Systems, **Jour. of Optimization Theory and Applications,** Vol. 50, 1986.

〈24〉 **Li, Xunjing and Chow, Shui-Nee,** Maximum Principle of Optimal Control for Functional Differential Systems, **Jour. of Optimization Theory and Applications**, Vol. 54, 1987.

〈25〉 **Li, Xunjing and Yao, Yunlong,** On Optimal Control for Distributed Parameter Systems, **Proc. of 8th IFAC World Congress,** Kyoto, Japan, 1981.

〈26〉 **Li, Xunjing and Yao, Yunlong,** Maximum Principle of Distributed Parameter Systems with Time-lags, in **Lecture Notes in**

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, А. А.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, Т. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〔27〕 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〔28〕 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〔29〕 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〔30〕 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〔31〕 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, А. А.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs complètment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs complètment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.AHCCCP,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs complètment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.**, Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.**, On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.**, Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control**, Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao**, The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia**, 4(3), 1984.

〈31〉 **You, Yuncheng**, Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.**, 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs complètment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs completment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.

Control and Information Sciences. No. 75, Springer-Verlag, 1985.

〈27〉 **Ляпунов, A. A.,** Sur les fonctions-vecteurs complètment additive, **Изв.АНСССР,серия матем.**, T. 4, 1940.

〈28〉 **McShane, E. J.,** On Multipliers for Lagrange Problems, **Amer. J. Math.** Vol.. 61, 1939.

〈29〉 **Russell, D. L.,** Penalty Functions and Bounded Phase Coordinate Control, **SIAM J. Control,** Vol. 2, 1965.

〈30〉 **Tsien, Jian and Chen, Zuhao,** The Convergence Rate of Penalty Function Methods for Optimal Control Problem, **Acta Mathematica Scientia,** 4(3), 1984.

〈31〉 **You, Yuncheng,** Optimal Control for Linear Systems with Quadratic Indefinite Criterion on Hilbert Spaces, **Chinese Ann. of Math.,** 4B(1), 1983.